



HAL
open science

Collaboration entre un humain, un robot et un système ambiant pour l'évaluation de comportements

Emmanuel Dumont

► **To cite this version:**

Emmanuel Dumont. Collaboration entre un humain, un robot et un système ambiant pour l'évaluation de comportements. Robotique [cs.RO]. Université de Technologie de Compiègne, 2019. Français. NNT : 2019COMP2484 . tel-02448718

HAL Id: tel-02448718

<https://theses.hal.science/tel-02448718>

Submitted on 22 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ TECHNOLOGIQUE DE COMPIÈGNE

Spécialité

Bio-ingénierie et Sciences et Technologies de l'Information et des Systèmes

Collaboration entre un Humain, un Robot et un Système ambiant pour l'évaluation de comportements

Auteur :

Emmanuel DUMONT

7 mai 2019

Directeurs de thèse :

M. Dan ISTRATE Co-directeur

Enseignant-Chercheur *Université de Technologie de Compiègne*

M. Mohamed CHETOUANI Co-directeur

Professeur des Universités *Université Pierre et Marie Curie*

Membres du jury :

M. Yacine AMIRAT Rapporteur

Professeur des universités *Institut Universitaire de Technologie de Créteil-Vitry*

M. Dominique VAUFREYDAZ Rapporteur

Maître de conférences *Université Grenoble Alpes*

Mme Laurence CHABY Examinatrice

Maître de conférences *Université Paris Descartes*

Mme Séverine DUBUISSON Examinatrice

Maître de Conférences *Aix-Marseille Université*

Mme Indira THOUVENIN Examinatrice

Enseignant-Chercheur *Université Technologique de Compiègne*



- Oh ! j'ai fait mieux depuis.

Cyrano de Bergerac, E.Rostand

Remerciements

Il est toujours compliqué de commencer la page d'un manuscrit qui sera très certainement la plus lue. Le bon mot doit s'appliquer ici tout autant que dans la conclusion d'un chapitre. J'espère que ces quelques tournures, *sans cire*, rendront état du fond de ma pensée.

Si Gandalf était seul pour lancer à l'aventure un hobbit, Dan Istrate et Mohamed Chetouani se sont complétés pour me faire sortir de ma comté dans le seul dessein de me faire gravir la montagne du Destin. Avec des discussions riches en conseils et recommandations, ils ont su m'épauler et m'instruire tout au long de ce voyage. Merci messieurs, pour votre patience, votre savoir et je tiens à vous présenter mes excuses pour les quelques cheveux blancs émergeant par ma faute. Gandalf n'est-il pas surnommé "Le Blanc" à la fin de l'histoire ?

Mes investigations m'auront amené à visiter l'INSEAD-Sorbonne Université Behavioural lab et l'hôpital Charles Foix pour récupérer les indices nécessaires et avancer dans les actes de ma thèse. Merci à Huang, au professeur Belmin et leurs équipes de m'avoir ouvert la porte en me donnant toutes les ressources pour bien conclure cette campagne.

Comme dirait Cyrano, "*Empanaché d'indépendance et de franchise,...*", je veux adresser ma reconnaissance envers ma famille. On lira souvent des phrases copiées-collées indiquant que "*si chaque famille ressemblait à la mienne, le monde serait en meilleur état*", de mon côté ? Le monde serait dans un chaos indescriptible...

En revanche, l'amour et le soutien seraient inconditionnels et illimités. A mes parents, mes grand-parents, mes soeurs, mes oncles et tantes, mon beau frère, mon neveu et ma nièce, merci. Puissent ces quelques mots, numériques ou imprimés, vous atteindre où que vous soyez.

Harry Potter, a construit une seconde famille autour de ses amis. C'est avec une grande émotion que j'écris mes remerciements envers les miens. Ils ont su, par un simple message ou le partage d'une boisson pétillante autour d'un jeu de cartes, me rappeler qu'un autre monde existait. Vous êtes si nombreux qu'il me faudrait des pages entières pour vous adresser individuellement les mots adéquats. Je me contenterai d'énoncer vos prénoms, comprenez que j'ai bien plus à dire.

Merci à Alexandre T, Axel, Benjamin, Cédric, Clément R, Damien, Estelle, Frédéric, Marianne, Marie-Alexandra, Marie-Sophie, Valentine ; Alexandre G, Ervan, Luc, Laurence et Mauro, Patrick ; Antonin, Céline, Laurianne ; Clément D, Nathalie, Pierre C ; Alex, Ao, Dimitri, Eliza, Iluvatar, Mr Patate, Nostal, Shep, Le Taulier, Titus.

A mes compagnons navigateurs et explorateurs (parfois de galères). Nous partîmes nombreux de la 302 et arrivâmes peu en pyramide, mais chacun en découvrant sa voie. À Anis, Alix, Caroline, Giovanna, Giuseppe, Hanan, Hakim, Isabelle, Laurence, Nathavy, Victor et Wail, que les discussions soient scientifiques ou non, merci pour les rires et les soutiens quotidiens. Un remerciement spécial est fait au Docteur Hajaoui qui, de sa parole, illuminait des après-midi trop courts à mon goût.

A toi, Pauline, par pudeur d'être observé par chaque lecteur, j'écris ces humbles mots sans même qu'ils osent venir de moi :

"*Et dès que je l'aperçois, Alors je sens en moi, Mon coeur qui bat*" Edith Piaf, *La Vie en Rose* - version réécrite par Yves Montand

Résumé

Évaluer un comportement humain c'est évaluer tous les marqueurs traduisant ce comportement (gestes, paroles interactions, etc.). L'observation par un humain de certains marqueurs tels que les expressions faciales, la prosodie ou encore la linguistique, nécessite une formation spécialisée. Pour faciliter l'évaluation du comportement, des échelles indiquant les observations à mener et les conclusions à faire sont employées. Ainsi, automatiser l'évaluation du comportement revient à automatiser l'analyse d'un environnement par le biais de plusieurs capteurs, puis analyser les signaux obtenus afin d'en extraire les marqueurs permettant la déduction du comportement observé.

Suite à la variabilité des observations de l'humain lors d'analyses trop spécifiques, de plus en plus d'études emploient ces systèmes automatiques d'observation et d'évaluation du comportement. L'objectif est d'assister l'analyse et l'évaluation humaine en exploitant des systèmes automatiques capables d'extraire des informations difficilement observables pour l'humain. En conséquence, la collaboration entre l'humain et les systèmes informatiques permet d'analyser plus d'éléments du comportement de manière fiable et objective.

Cette thèse propose une approche de l'analyse du comportement s'appuyant sur la collaboration entre l'humain et un système automatique. Nous avons mis en place une plate-forme électronique et informatique composée d'un robot mobile et d'un système ambiant afin d'évaluer le comportement humain. Cette plate-forme se définit comme étant :

- Modulaire à l'ajout ou le retrait de capteurs : L'ajout et la suppression de capteurs est faisable sans qu'un système ne soit impacté autrement que sur ses performances à reconnaître précisément les comportements ;
- Accessible à la lecture des données enregistrées : L'utilisation d'ontologies, en tant que base de données sémantiques et logiques, rend la plate-forme utilisable et accessible aux personnes non familiarisées aux systèmes informatiques complexes ;
- Robuste aux ambiguïtés : Chaque système de la plate-forme (ambiant ou robot) est indépendant et a sa propre représentation de l'environnement. Cependant, ils collaborent entre eux pour répondre aux incohérences ou aux manques d'informations durant l'accomplissement d'une tâche.

A partir de la plate-forme présentée précédemment, nous analysons et mesurons la qualité de l'interaction entre un patient et un soignant lors d'une prise de sang réalisée en conditions habituelles. Pour cela, nous utilisons deux méthodes de renseignement des échelles : par un observateur présent lors du soin et par une étude de l'enregistrement vidéo réalisé durant le soin par la plate-forme. Nous émettons l'hypothèse que la présence d'un système automatique d'aide au diagnostic lors de l'analyse des vidéos enregistrées limite la complexité de l'évaluation du comportement et améliore l'objectivité de l'analyse.

Table des matières

1	Introduction	13
1.1	L’observation du comportement humain	13
1.2	Les systèmes automatiques d’observation du comportement humain	15
1.2.1	La fusion de données	15
1.2.2	Ontologie	16
1.3	Applications	18
1.3.1	Les Activités de la Vie Quotidienne	18
1.3.2	Le comportement du patient et du soignant lors d’un soin	20
1.4	Objectifs et contributions	21
1.4.1	Considérations éthiques pour les expériences et les données numériques	23
1.5	Organisation du document	25
2	Une plate-forme de modélisation du comportement	27
2.1	Description de la plate-forme	27
2.1.1	Un système de systèmes	27
2.1.2	Logiciels et algorithmes employés	29
2.2	Processus de collaboration inter-système	29
2.2.1	Description du processus	29
2.2.2	Aperçu du processus à travers un scénario	30
2.2.3	Spécifications spatiales au sein de l’ontologie	31
2.3	Choix techniques	32
2.4	Conclusion	33
3	Collaboration entre un robot et un système ambiant pour la reconnaissance d’activités	35
3.1	Étude sur la reconnaissance des activités de la vie quotidienne	35
3.1.1	Objectif et hypothèse de l’étude	35
3.1.2	Schématisation de l’observation des activités	36
3.1.3	Caractéristiques techniques et algorithmique	38
3.2	Évaluation des résultats et observations	41
3.3	Conclusion de l’étude et discussion	44
4	Évaluation automatique de l’interaction Patient/Soignant au cours d’un soin	45
4.1	Objectif, méthodologie et matériel de l’étude	45
4.1.1	Objectif, hypothèses et définitions	45
4.1.2	Une étude, deux expériences	46
4.1.3	Spécifications techniques	46
4.2	Expérience réalisée au sein de l’INSEAD Behavioural Lab	51
4.2.1	Protocole	51
4.2.2	Les évaluations	52

4.2.3	Résultats	53
4.2.4	Conclusion	55
4.3	Expérience réalisée au sein de l'hôpital Charles Foix	56
4.3.1	Protocole	58
4.3.2	Problématiques environnementales et adaptation du protocole initial	60
4.3.3	Analyse des résultats	61
4.3.4	Conclusion	66
5	Conclusions et perspectives	67
5.1	Perspectives	69
A	ECPAI	77
B	ECIIPA	79
C	Résultats comparatifs des échelles patient/soignant de l'interaction simulée	81
D	Lettre de compte rendu de l'étude	87
D.1	Explication de ce type de recherche	87
D.2	Description plus détaillée de l'étude	87
D.3	A propos du Centre Sorbonne Universités-INSEAD	88
D.4	A propos de Sorbonne Universités	88
D.5	A propos de l'INSEAD	88
D.6	Le chercheur	88
E	LETTRE D'INFORMATIONS ADRESSEE AUX PATIENTS POUR L'ÉTUDE	89
E.1	BUT DE L'ETUDE	89
E.2	BENEFICES ATTENDUS	89
E.3	DEROULEMENT DE L'ETUDE	89
E.4	RISQUES POTENTIELS	90
E.5	FRAIS DE L'ETUDE	90
E.6	LEGISLATION - CONFIDENTIALITE	90
F	LETTRE D'INFORMATIONS ADRESSÉE AUX SOIGNANTS POUR L'ÉTUDE	93
F.1	BUT DE L'ETUDE	93
F.2	BENEFICES ATTENDUS	93
F.3	DEROULEMENT DE L'ETUDE	93
F.4	RISQUES POTENTIELS	94
F.5	FRAIS DE L'ETUDE	94
F.6	LEGISLATION - CONFIDENTIALITE	94

Table des figures

1.1	Exemple de marqueurs annotés par un observateur humain remplissant des échelles d'évaluation du comportement	14
1.2	Architecture de traitement d'informations récupérées par des capteurs	16
1.3	Schéma du processus de collaboration entre la plate-forme et un observateur humain	21
2.1	Architecture électronique et informatique de la plate-forme	28
2.2	Schéma du processus de collaboration inter-système pour de la recherche d'objet	30
2.3	Robot mobile oculus prime	33
3.1	Schéma de l'étude. L'Observateur Référent (OR) annote les activités en cours. L'observateur Participant (OP) utilise les informations extraites par la plate-forme afin de reconnaître et annoter les activités observées.	36
3.2	a) Aperçu de la vue globale des OR ; b) Aperçu des informations que les OP ont pour évaluer les activités	37
3.3	Séquencement des horaires des activités observées par les observateurs référents	38
3.4	Capteurs employés dans le système ambiant : a) Kinect, b) Capteur de Sommeil, c) Capteur de mouvement infrarouge	39
3.5	Disposition du système ambiant dans l'environnement et séquencement des activités. 1) <i>Le participant dort</i> 2) <i>Après s'être levé, le participant marche jusqu'à l'armoire où il va s'habiller</i> 3) <i>Le participant quitte la chambre et marche jusqu'à la cuisine pour préparer un repas</i> 4) <i>Le participant rejoint la table de la salle à manger, s'installe, mange en regardant la télévision</i> 5) <i>Le téléphone du participant sonne, il répond en quittant la table</i>	40
3.6	Les capteurs utilisés et les propriétés d'objets auxquels ils se rattachent pour nourrir l'ontologie	40
3.7	Séquencement des activités d'après les observateurs référents et les observateurs participants	41
3.8	Exemple d'application de l'équation 3.1 pour l'activité S'Habiller évaluée par un observateur référent et un observateur participant	42
4.1	Schéma explicatif de la première expérience	48
4.2	Point de vue des différents kinects et de la caméra grand angle	48
4.3	Séquencement des activités d'après les observateurs référents et les observateurs participants	49
4.4	Photographie illustrant la mise en place du chariot médical. La kinect positionnée sur le plateau principal du chariot permet de filmer le soignant lorsqu'il rentre et sort de la chambre. En réorientant le chariot durant le soin, la kinect filme l'interaction. Les tiroirs et le plateau peuvent être utiliser pour poser du matériel de soin	50

4.5	Exemple de comparaison d'évaluation du comportement d'un soignant réalisée par un observateur référent et la plate-forme	53
4.6	Schéma des évaluations des comportements du patient et du soignant au cours de cette étude. L'interaction patient-soignant est évaluée de trois manières différentes. La première est réalisée par un observateur présent dans la chambre (OR) durant le soin. La deuxième analyse est réalisée par des observateurs utilisant les enregistrements bruts de la plate-forme (OV). Enfin, la dernière analyse est réalisée à posteriori par un observateur (OS) s'appuyant sur les informations extraites par le système automatique	57
4.7	1) Le chariot médical comprenant le système embarqué est préparé en avance ; 2) Le soignant entre dans la chambre, la caméra le filme durant la première interaction avec le patient 3) le soignant positionne le chariot de manière à ce que la caméra puisse enregistrer la prise de sang 4) Le soignant sort avec le chariot et quitte la chambre. OR représente l'Observateur référent dans la chambre tout au long de l'enregistrement. Il est chargé d'évaluer l'interaction Patient/Soignant.	59
4.8	Prise de vue du système lors d'un soin. Dans l'image de gauche, le patient est sous les draps. Dans celle de droite, la luminosité et le positionnement du patient n'ont pas permis de réaliser les mêmes traitement d'information que dans l'étude précédente faite dans un environnement contrôlé.	60
4.9	Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d'observateurs évaluant le comportement de la même manière : Lorsque les deux observateurs ont évalué de manière similaire, la case correspondante dans l'échelle est représentée en vert, en rouge sinon	62
4.10	Résultats de la comparaison des échelles annotées par les observateurs vidéos	62
4.11	Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d'observateurs évaluant le comportement de la même manière : Les cases vertes correspondent à l'évaluation de l'observateur référent, les cases rouges sont les évaluations discordantes faites par les observateurs vidéos	64
4.12	Résultats de la comparaison des échelles annotées par les observateurs vidéos et l'observateur référent	64
4.13	Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d'observateurs évaluant le comportement de la même manière : Les cases vertes correspondent à l'évaluation de l'observateur référent, les cases rouges sont les évaluations discordantes faites par l'observateur système	65
4.14	Résultats de la comparaison des échelles annotées par l'OR et l'OS	65
C.1	Ensemble des échelles complétées par les patients et les soignants. En vert, la correspondance des évaluations, en rouge la discordance	82
C.2	Ensemble des échelles complétées par les soignants et l'observateur vidéo. En vert, la correspondance des évaluations, en rouge la discordance	83
C.3	Ensemble des échelles complétées par les patients et l'observateur vidéo. En vert, la correspondance des évaluations, en rouge la discordance	84
C.4	Ensemble des échelles complétées par le système autonome	85
C.5	Ensemble des échelles complétées par le système (en vert) comparées à celles de l'observateur vidéo (en rouge)	86

Liste des tableaux

2.1	Tableau des algorithmes employés et leur référence	29
3.1	Tableau des propriétés chaînées utilisées au sein de l'ontologie	39
3.2	Tableau des abstentions et informations des évaluations des observateurs participants	43
4.1	Résultats de la comparaison des échelles annotées par les observateurs vidéos et l'observateur référent	54
4.2	Résultats de la comparaison des échelles annotées par l'OR et le patient	54
4.3	Résultats de la comparaison des échelles annotées par l'OR et le soignant	54
4.4	Résultats de la comparaison des échelles annotées par l'OR et le système automatique	55

Chapitre 1

Introduction

Depuis quelques années se développent les concepts de reconnaissance automatique d'activités de la vie courante et reconnaissance automatique de comportements. Ces concepts s'articulent autour de domaines de recherche comme le "bien-vieillir à domicile" ou plus généralement des sciences du comportement humain. Les technologies alors mises en place utilisent un ensemble de capteurs positionnés de manière fixe dans des salles (salon, cuisine, chambre hospitalière, bureaux, etc.) ou embarqués sur des robots mobiles pouvant interagir, ou non, avec l'être humain. Un système informatique récolte et analyse ensuite les données afin de déduire une activité ou un comportement correspondant aux modèles définis au sein du système.

Les plate-formes mises en place se sont peu à peu complexifiées pour répondre aux problématiques d'environnements contraignants ou d'activités et comportements difficiles à reconnaître. L'émergence de systèmes de systèmes [Jamshidi, 2008b] ; *définissables comme des systèmes complexes composés eux mêmes de plusieurs sous-systèmes complexes* ; a entraîné l'apparition de nouvelles problématiques sur la fusion de données multimodales, la collaboration des systèmes et plus généralement sur le concept "d'intelligence ambiante" ; *terme désignant les environnements électroniques dont les activités de traitement sont centrées autour de l'être humain* [Wright et al., 2008]. Par exemple, au sein d'une même plate-forme, chaque système est composé de plusieurs capteurs, chacun donnant une ou plusieurs informations complémentaires sur l'environnement. Une première question se pose lorsque les données obtenues sont disparates. Dans l'optique d'une collaboration entre les sous-systèmes composant une plate-forme d'intelligence ambiante, une autre problématique est celle du désaccord entre eux.

Dans les sections qui suivent, nous présentons les problématiques liées à l'observation des activités de la vie courante et des comportements humains ainsi que les axes de recherches des systèmes de systèmes. Nous finirons ce premier chapitre avec une description de l'organisation du document.

1.1 L'observation du comportement humain

L'analyse d'un comportement permet d'établir une description générale des mécanismes qui le produisent [Baer et al., 1968]. Observer un comportement humain consiste à annoter et séquencer l'ensemble de ces mécanismes : mouvements, actions, paroles et expressions réalisés par un être humain en réaction à un évènement ou stimuli (voir figure 1.1). L'évaluation du comportement qui consiste à mesurer et quantifier les réactions observées, leur fréquence, leur magnitude (ou intensité) et la durée, ainsi que les applications qui lui sont liées, permet de déceler et comprendre des maladies [Iranzo et al., 2006, Bourgeois, 1990], des troubles du comportements [Lord et al., 1989, Virués-Ortega, 2010] ou encore renforcer

des comportements [Shabani and Fisher, 2006] et des méthodes d'éducation [Barrish et al., 1969, Cooper, 1982, Stocco and Thompson, 2015].

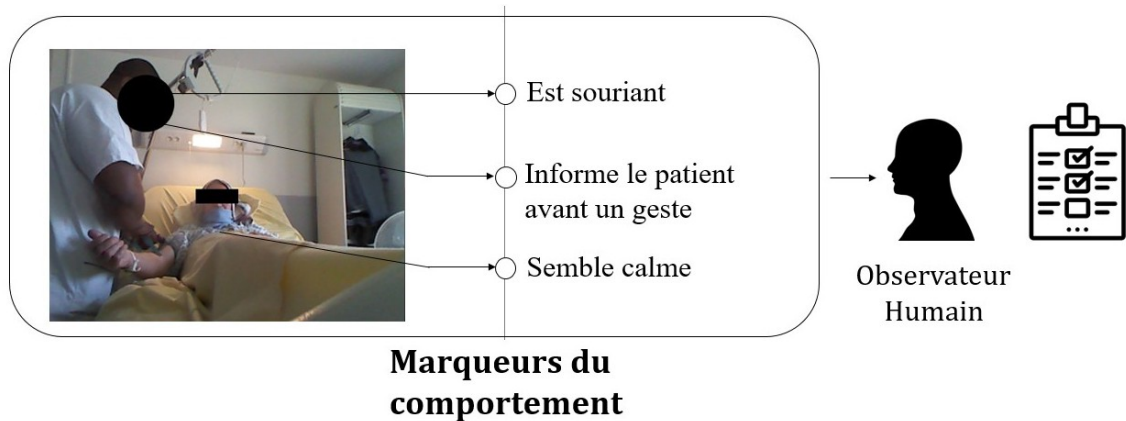


FIGURE 1.1 – Exemple de marqueurs annotés par un observateur humain remplissant des échelles d'évaluation du comportement

Afin d'observer et mettre en évidence un comportement dans un contexte particulier différentes approches existent [Altmann, 1974, O'Neill et al., 2015] :

1. Le modèle antécédent-comportement-conséquence (modèle ABC - Antecedent Behavior Consequences) : Cette analyse consiste à observer une séquence temporelle où les antécédents d'un comportement, le comportement et les conséquences surviennent immédiatement après le comportement. Cette approche permet de comprendre l'évènement ou le stimuli déclenchant le comportement et ses conséquences [Skinner, 1966].
2. La méthode d'enregistrement des intervalles : Cette méthode consiste à analyser uniquement les périodes de temps où le comportement a lieu pour l'évaluer.
3. Les analyses fréquentielles ou événementielles dénombrent la manifestation d'un comportement sur une période temporelle. Ces analyses permettent de comparer des comportements d'individus entre eux ou encore le taux d'occurrences.

Lors de l'observation, l'analyse ou l'évaluation d'un comportement, un ou plusieurs observateurs sont présents afin d'annoter les comportements tels qu'ils les perçoivent. L'utilisation d'échelles par ces observateurs est courante afin d'annoter des informations comme le début, la fin d'un évènement ou encore mesurer les réactions. Trois rôles sont possibles pour chaque observateur :

1. Un observateur présent dans la situation qui interagit tout en observant le comportement,
2. Un observateur extérieur qui observe la situation en dehors de la scène (vidéo, glace sans tain, etc.),
3. Un observateur qui s'auto-évalue et répond lui même à un questionnaire.

Cette dernière méthode est particulière puisqu'elle compte sur sa propre perception de soi et des évènements pouvant ainsi créer un biais dans les informations [Stone et al., 1999, Aigrain et al., 2016].

1.2 Les systèmes automatiques d'observation du comportement humain

Au cours des dernières années, de nombreux efforts ont été réalisés pour automatiser l'analyse du comportement par des systèmes [Schaal, 1999, Atkeson et al., 2000]. Un système est défini comme un ensemble technologique capable de percevoir l'environnement et les êtres humains qui s'y situent afin d'analyser leurs comportements.

Des domaines de recherche pluridisciplinaires étudient et conçoivent des systèmes capables d'analyser ou simuler des états affectifs humains (on parle alors d'*affective computing*) [Picard et al., 1995]; ou des robots sociaux [Kanda et al., 2004, Fong et al., 2003]. Ces derniers sont définis comme étant "capables de reconnaître d'autres robots ou humains et d'engager une interaction sociale avec" [Billard and Dautenhahn, 1997, Breazeal, 2004].

Une des premières application de la robotique sociale concerne l'assistance aux personnes à travers les interactions sociales. Une partie de l'application donnée à ces robots concerne le domaine de la santé. Par exemple dans le domaine de l'autisme [Scassellati et al., 2012, Anzalone et al., 2014], des personnes âgées [Broekens et al., 2009, Wada et al., 2005] ou de la télémédecine [Tsai et al., 2007, Chan et al., 2008].

1.2.1 La fusion de données

Chaque type de système s'emploie à récolter des informations de l'environnement dans lequel il est placé pour ensuite les traiter. Suite à ces traitements, les données déduites sont stockées en attendant leur utilisation (voir la figure 1.2).

Les données extraites d'un environnement dépendent de l'étude d'un phénomène (détection d'un élément spécifique, reconnaissance d'un comportement, etc.). Les capteurs et les algorithmes sont sélectionnés et implémentés au sein du système à mettre en place en fonction des contraintes liées à cette étude. Il est rare qu'une seule donnée permette la représentation complète du phénomène d'intérêt. En conséquence, obtenir plusieurs données complémentaires ou redondantes à analyser est nécessaire pour compléter le point de vue global du système sur le phénomène. Une fois récoltées, ces données sont à comparer et assembler entre elles. Cette étape est celle de la fusion des données (voir la figure 1.2).

Plus le nombre de données est important, plus la fusion de données est complexe [Khalghi et al., 2013]. Ceci est dû à la diversité, le nombre et le type des données [McGurk and MacDonald, 1976, Turk, 2014]. En fonction des études, plusieurs méthodes de fusion de données sont employées et comparées entre elles afin de déterminer la plus adéquat. Par exemple [Gan and Harris, 2001] compare l'application d'un filtre de Kalman sur les vecteurs de données multimodales mesurées [Kalman, 1960] à l'estimation de l'erreur quadratique moyenne minimale de ces mêmes données mesurées. L'hypothèse que Gan cherche à démontrer est que les résultats obtenus sont semblables. Cependant le temps de calcul de la seconde méthode étant sensiblement inférieur, elle serait à privilégier en fonction des applications.

Un travail similaire a été fait par [Roecker and McGillem, 1988]. Deux méthodes de fusion de données basées sur les filtres de Kalman sont comparées; la moyenne de la covariance du vecteur d'état et la méthode de compression décrite par Willner dans [Willner et al., 1976]; pour suivre la trace d'un élément mobile en utilisant plusieurs capteurs. L'objectif étant d'améliorer et d'optimiser les résultats obtenus par diverses sources.

Ainsi, chaque système cherchant à étudier un phénomène à l'aide de capteurs spécifiques met en place une méthode de fusion de données inspirées d'approches mathématiques; par

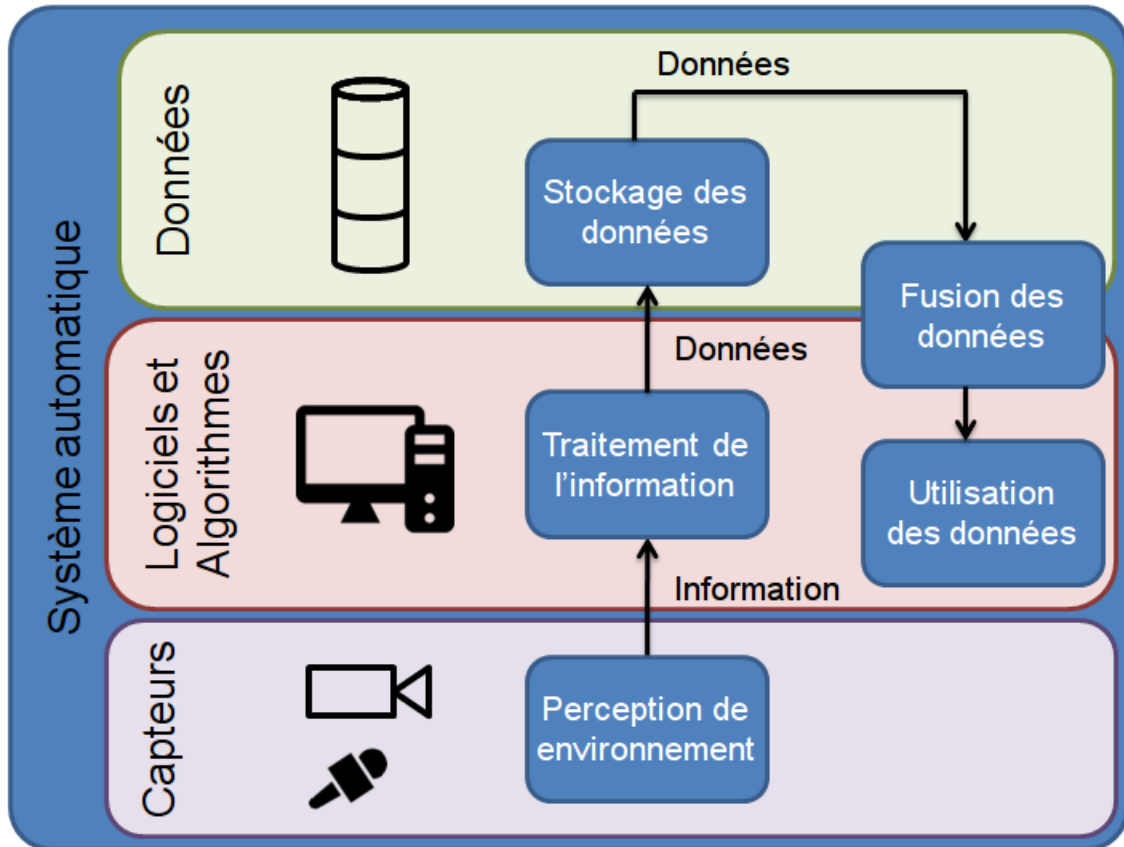


FIGURE 1.2 – Architecture de traitement d'informations récupérées par des capteurs

exemple les filtres de Kalman ou encore la théorie de l'évidence de Dempster-Shafer [Shafer, 1976]. Cette dernière est utilisée par [Hong et al., 2009] dans le cadre de données multimodales provenant d'un appartement intelligent. L'objectif étant d'automatiser la reconnaissance d'activités de la vie quotidienne.

Une autre partie de la recherche sur la fusion automatique s'axe autour de la formalisation des données extraites [Khaleghi et al., 2013]. L'objectif est de fusionner des données préalablement traduites dans un langage formel en utilisant des concepts standards et logiques mis en place au sein d'ontologies [Nowak, 2003, Lemaignan, 2013]. Outre la standardisation des données multimodales, ces systèmes ont l'avantage de rendre accessible les données à n'importe quel utilisateur puisqu'ils s'appuient sur une syntaxe proche du langage courant.

1.2.2 Ontologie

Les ontologies sont utilisées dans plusieurs domaines : Militaire [Bowman et al., 2001], Santé [Charlet et al., 2009], Agriculture [Roussey et al., 2011], etc.

Au sein du domaine médical, les ontologies sont utilisées dans deux types d'applications [Charlet et al., 2009] : la première vise à la réalisation d'outils logiciels représentant de manière formelle les connaissances médicales et déterminants les processus d'inférences de diagnostics. La seconde repose aussi sur une sémantique forte mais la combine avec des ressources linguistiques pour n'automatiser que certains traitements d'informations prove-

nant d'un utilisateur, par exemple lors du dialogue entre un patient et un soignant.

En informatique, une ontologie permet de nommer et définir les classes (Objet, Humain, Capteur, etc.), les attributs (propriétés temporelles, spatiales et physiques), les relations et les événements (changements d'attributs ou de relation) d'un concept qui existe au sein d'un domaine pouvant être représenté formellement. Autrement dit, une ontologie représente et normalise le monde physique avec le langage naturel.

Dans le cadre des systèmes robotiques, elles sont utilisées pour normaliser les informations provenant de sources multimodales facilitant leur comparaison. Les ontologies stockent des informations de manière sémantique et sont donc accessibles à la lecture pour un humain [Chandrasekaran et al., 1999]. Elles fonctionnent alors comme une base de données sémantiques et logiques. Un raisonneur *raisonne* sur les données enregistrées en leur appliquant les propriétés et chaînes de propriétés qui y sont décrites afin d'en déduire de nouvelles informations [Sirin et al., 2007, Carroll et al., 2004]. Au final, l'ensemble des données et des propriétés présentes au sein de l'ontologie permettent une représentation de l'environnement et du contexte.

Une ontologie stocke des informations en utilisant le format RDF (Resource Description Framework). Ce format est un modèle standard pour l'échange de données sur le web. Le principe consiste à décrire des connaissances sous forme d'assertion ou "triplet" comprenant le sujet, un objet et le lien entre les deux, par exemple "*Ela chante Summertime*", où "*Ela*" est le sujet caractérisé comme appartenant à la classe "*Humain*", "*Summertime*" est un objet caractérisé comme appartenant à la classe "*Chanson*". La propriété *chante* est la relation entre le sujet et son objet.

Une propriété est comparée à un verbe conjugué : *chante*, *cuisineDans* ou encore *estProcheDe*. Ces propriétés peuvent ne s'appliquer qu'à des classes en particulier. Ainsi, les classes que lie une propriété peuvent être spécifiées en tant que *domain* et *range*, respectivement classe du sujet et classe des objets, de manière à sécuriser la représentation de l'environnement. Si une assertion est ajoutée à l'ontologie telle qu'une des classes du triplet ne correspond pas ni au *domain* ni à la *range* le raisonneur détectera une incohérence et empêchera l'ajout de ce triplet dans l'ontologie.

Une chaîne de propriétés consiste à créer une nouvelle propriété qui déduira une information à partir des propriétés précédemment définies. Par exemple, la propriété chaînée liant *estActivéDans* et *estProcheDe* met en place la propriété *cuisineDans*, ainsi les triplets d'informations suivants [*four estActivéDans cuisine*] et [*Martin estProcheDe four*] impliquent la propriété chaînée [*Martin cuisineDans cuisine*]. Cette dernière est automatiquement déduite par le raisonneur de l'ontologie lorsque les deux triplets [*four estActivéDans cuisine*] et [*Martin estProcheDe four*] sont ajoutés au sein de l'ontologie.

Dans ces exemples, les activités de la vie quotidienne en elles mêmes sont des données de plus haut niveau. Elles sont établies comme un triplet : [*Humain faitActivitéDans Lieux*] au sein de l'ontologie et sont déduites d'autres informations collectées par les capteurs en utilisant des propriétés chaînées.

Une propriété d'objet est dite symétrique : "Pour chaque individu x , connecté à y par une propriété symétrique, y est alors connecté à x par cette même propriété"¹. Ce qui signifie que suite à la déclaration dans l'ontologie : "*Martin estProcheDe Jean*", le raisonneur créera automatiquement la déclaration : "*Jean estProcheDe Martin*".

Une propriété d'objet est fonctionnelle : "Pour chaque individu x , il y a au moins un indi-

¹ Voir *Symmetric Object Properties* https://www.w3.org/TR/owl2-syntax/#Symmetric_Object_Properties

vidu y distinct tel que x est connecté par *estDans* à y "². En d'autres mots, une première déclaration "*Martin estDans cuisine*" suivi d'une seconde déclaration "*Martin estDans chambre*" va créer une incohérence suite au fait que *Martin* ne peut être uniquement que dans une *Pièce* à la fois. Au sein de l'ontologie, cela se traduit par une seule déclaration "*Martin estDans [Pièce]*", peu importe la pièce.

En définitive, une donnée formalisée au sein d'une ontologie peut être comparée ou complétée avec n'importe quelle autre donnée formalisée au sein de la même ontologie et ce, peu importe leur provenance. On parle alors de données de haut niveau ou encore de connaissance. Ainsi, une information spécifique provenant d'un type de capteur sur un robot peut être comparée à une information fondamentalement identique provenant d'un capteur de type différent.

Dès lors, les ontologies peuvent être employées pour fusionner les connaissances de différents systèmes intégrés dans des plate-formes complexes [Madni and Sievers, 2014] en garantissant l'interopérabilité entre chaque système. Ces plate-formes, aussi appelées systèmes de systèmes [Jamshidi, 2008a], peuvent être composées d'architectures comprenant plusieurs robots connectés entre eux (aussi appelé *cloud robotics* [Kuffner, 2010]), ou la combinaison de systèmes robotiques et de systèmes ambiants.

Ces derniers sont définis comme l'ensemble des capteurs placés dans un lieux pour en récupérer les informations environnementales. Ce terme est un dérivé de l'intelligence ambiante (AmI) qui est l'environnement électronique mis en place pour aider de manière proactive l'être humain dans sa vie courante [Ramos et al., 2008]. Les paradigmes de l'AmI se construisent principalement autour des points suivants : Les systèmes embarqués employés au sein de l'AmI doivent s'intégrer dans l'environnement avec les contraintes qui lui sont propres. Par exemple, les travaux de A. Dunkels [Dunkels et al., 2004] proposent une architecture électronique et informatique optimisée, entre autre en terme de consommation énergétique, pour des capteurs environnementaux.

Le deuxième, l'adaptabilité et la personnalisation de ces systèmes à l'environnement. Un exemple concerne la planification des actions du système en fonction de l'utilisateur [Amigoni et al., 2005]. Cette adaptabilité comprend aussi la notion de modularité permettant de modifier l'état du système durant son utilisation pour l'assortir aux besoins.

1.3 Applications

Cette thèse s'inscrit dans le cadre des projets de la chaire EBiomed intégrée au laboratoire BioMédical et BioIngénierie - BMBI UMR CNRS 7338 et l'Institut des Systèmes Intelligents et de Robotiques - ISIR CNRS UMR 7222. Dans le cadre des travaux de la chaire et des partenariats avec les Hôpitaux Universitaires de la Pitié Salpêtrière Charles-Foix, nous nous sommes intéressés de près à l'observation des activités de la vie quotidienne chez les personnes âgées et à l'évaluation de l'interaction entre un patient et un soignant lors d'un soin.

1.3.1 Les Activités de la Vie Quotidienne

Les AVQ sont l'ensemble des activités essentielle pour un humain pour vivre au quotidien et s'occuper de soi. L'association américaine des ergothérapeutes (AOTA) [Roley et al., 2008] les résumes comme étant :

² Voir *Functional Object Properties* https://www.w3.org/TR/owl2-syntax/#Functional_Object_Properties

- Se laver en prenant une douche ou un bain
- Le contrôle des sphincters
- S’habiller
- Manger (déglutir, amener la nourriture de son assiette à sa bouche, etc.)
- L’hygiène et soins personnels
- La toilette
- La mobilité fonctionnelle (se déplacer au quotidien)

Les Activités Instrumentales de la Vie Quotidienne (AIVQ) sont des activités qui ne sont pas fondamentales pour vivre. En revanche, être en mesure de les réaliser permet de vivre de manière autonome. L’AOTA en dénombre douze :

- Prendre soin des autres,
- Prendre soin des animaux de compagnie,
- Élever ses enfants,
- Communiquer,
- Conduire ou utiliser les moyens de transports publics,
- Gérer les finances,
- Gérer sa santé,
- Entretien son domicile,
- Cuisiner incluant aussi le nettoyage,
- Participer à des activités religieuses,
- Suivre des procédures de sécurité et d’urgences,
- Faire des courses.

Dans le milieu médical, évaluer les capacités d’une personne à réaliser ces activités permet de prendre des décisions sur le patient comme le garder hospitalisé ou au contraire le laisser rejoindre son domicile. De plus, en fonction des besoins du patient, des aides à domicile peuvent lui être attribuées [[de la Solidarité et de la Santé web, 2018](#)]. Dans d’autres cas, afin d’aider la personne à redevenir indépendante dans ses activités, évaluer ses AVQ avant l’intervention d’un ergothérapeute, puis après, permet de s’assurer que les objectifs fixés sont correctement atteints.

Dans la section 3.1, nous nous intéresserons principalement aux activités suivantes *Cuisiner, Dormir, s’Habiller, Manger, Regarder la télévision et Téléphoner*. Notre travail ne consiste pas à analyser et reconnaître les activités en fonction de divers niveaux de complexité comme la déglutition pour l’activité *Manger*. Notre objectif à travers cette reconnaissance des activités est de pouvoir reconnaître et séquencer les AVQ durant une période de temps. Ainsi, *Cuisiner* sera symbolisée par l’utilisation d’un réfrigérateur et d’un micro-onde. *S’Habiller* consistera à ajouter plusieurs vêtements à une tenue complète initiale. Pour des raisons éthiques, les activités en lien avec la toilette et l’hygiène ne sont pas étudiées.

La robotique et les systèmes ambiants sont de plus en plus étudiés et impliqués dans l’amélioration du bien-être et des conditions de vies des personnes âgées [[Rashidi and Mihailidis, 2013](#), [Queirós et al., 2015](#)]. Ces systèmes sont utilisés pour effectuer des tâches d’observation traditionnellement réalisées par des humains. Dans le cadre des Activités de la Vie Quotidienne (AVQ), un ensemble de travaux récents s’attellent à automatiser cette reconnaissance d’activités (voir [[Lara et al., 2013](#)]). En effet, la reconnaissance d’AVQ implique qu’un référent humain soit présent dans l’environnement ou visionne des enregistrements. Outre une question éthique sur la visualisation de certaines activités, comme tout ce qui est en rapport à l’hygiène et la toilette, le référent reste monopolisé durant le

temps nécessaire pour assister ou visionner les activités.

Une solution est d’employer des systèmes automatiques [Tapia et al., 2004]-[Chen et al., 2012]. Cependant, ces systèmes sont soumis à des limitations ou des incertitudes induisant des erreurs dans leur analyses (angle mort dans l’angle de vision d’une caméra, informations contradictoires provenant de différents capteurs, etc.) entraînant des résultats de reconnaissance inférieurs à celles d’un référent humain. A l’inverse, limiter les capteurs ambiants et les remplacer par un système robotique mobile [Xia et al., 2015], ou combiner les deux systèmes [Hu et al., 2014], permet de limiter les problèmes de positionnement de ces capteurs, limiter les angles morts et donc la perte d’informations. De la même manière, augmenter le nombre de capteurs et les positionner sur un ensemble d’objets du quotidien (portes, bouilloires) permet d’améliorer les résultats de reconnaissance des activités [Tapia et al., 2004] et [Van Kasteren et al., 2008]. Cependant, la perte d’informations, par défaillance d’un capteur par exemple, limite ces systèmes. Enfin dans chacun d’eux, le remplacement d’un type de capteur par un autre (remplacement de caméra par des microphones par exemple) empêche le fonctionnement de la plate-forme mise en place.

1.3.2 Le comportement du patient et du soignant lors d’un soin

Dans cette thèse, nous avons concentré une partie de nos travaux sur l’évaluation de l’interaction entre le patient et le soignant lors d’un soin. A l’hôpital, la relation soignant-patient est primordiale dans la prise en charge hospitalière, le manque de temps pour les réaliser rendant les actes médicaux plus techniques que sociaux. Or, il a été montré que la qualité de la relation entre le soignant et le patient durant la prise en charge a un impact concret et direct sur la qualité du soin ressenti par le patient mais aussi par le soignant : [Hall et al., 1995] ont spécifié que le comportement non-verbal du soignant a une influence directe sur l’appréciation que lui porte son patient. Ainsi, afin de mieux former le personnel médical, l’importance de l’analyse du comportement du patient et/ou du soignant a été mise en évidence par [Mast, 2007], [Pawlikowska et al., 2012] et [Kurtz et al., 2016]. Dans la pratique clinique habituelle, il n’existe pas de moyen d’objectiver cette relation essentielle. Ce constat a motivé le développement d’échelles de mesures permettant d’objectiver l’interaction entre le patient et le soignant lors d’un acte médical :

- L’échelle d’Evaluation du Comportement de la Personne Aidée au cours de l’intervention avec un Intervenant – ECPAI, voir l’échelle ECPAI dans l’annexe A
- L’Echelle de Comportement de l’Intervenant au cours d’une Interaction avec une Personne Aidée - ECIIPA, voir l’échelle ECIIPA dans l’annexe B

Lors d’un soin, les comportements du patient et du soignant sont observés et annotés par un observateur, appartenant au personnel médical de l’hôpital, présent dans la salle lors du soin. [Girard and Cohn, 2016] précise que l’analyse de l’observateur peut varier au cours du temps; que le comportement des personnes observées peut changer face à une évaluation manifeste et que le coût de cette observation peut être important (formation, modélisation de personnelle, etc.).

[Olugbade et al., 2018], et les articles publiés lors de l’Audio Video Emotions Challenges (AVEC) [Valstar et al., 2013] et [Valstar et al., 2014], ont montré que l’analyse automatique du comportement permet de détecter et reconnaître des comportements que l’humain ne peut analyser avec certitude sans une formation importante.

L’analyse de comportements humain est très souvent employée dans le domaine médical (dépression, peurs et phobies, douleur chronique, déficit de l’attention, etc.). L’analyse peut se faire soit par les observations d’un personnel médical formé, soit avec l’auto-évaluation du patient par le biais de questionnaires. En effet, l’emploi de ces systèmes pour compléter

une évaluation permettrait de percevoir ce que l’humain ne peut déceler lorsqu’il évalue lui-même un comportement [Cummins et al., 2015]. Ainsi, dans le cadre du diagnostic de la dépression, le succès de ces analyses provient souvent de l’expérience du clinicien et des informations évidentes des patients exprimant leur motivation. [France et al., 2000], [Yingthawornsuk et al., 2007] et [Scherer et al., 2013] ont étudié des systèmes automatiques de classification pour évaluer automatiquement la prosodie et la voix et détecter les patients à tendances dépressives ou suicidaires. Les résultats démontrent que l’utilisation de systèmes automatiques d’aide au diagnostic limite la complexité de l’évaluation de ces patients.

Dans un autre cadre, les patients atteints de douleurs chroniques sont encouragés à conserver des activités physiques. Afin de passer la barrière psychologique de la douleur, ils sont encouragés à évaluer eux-mêmes la performance des mouvements qu’ils estiment pouvoir faire (Movement Related Self-Estimated - MRSE). Les cliniciens sont chargés d’évaluer leurs estimations avec l’activité réalisée. [Olugbade et al., 2018] démontre que chaque clinicien a sa propre méthode d’analyse et propose dans un premier temps de normaliser l’évaluation, puis de l’automatiser avec l’aide d’un système électronique et informatique. Les résultats obtenus indiquent que les évaluations faites par le système automatique sont similaires aux évaluations des cliniciens.

L’utilisation d’un système automatique apparaît donc comme une alternative intéressante afin d’assister l’observation des comportements soignant-patient par un observateur extérieur au soin.

1.4 Objectifs et contributions

Cette thèse, se situe dans les domaines de l’analyse automatique des activités de la vie quotidienne et l’évaluation du comportement dans le cadre d’une interaction entre humains. L’objectif principal est de mettre en place une plate-forme collaborative capable de répondre aux problématiques de fusion de données multimodales pour la collaboration entre les systèmes qui la compose. De plus, elle observe et analyse les comportements humains pour aider l’analyse des évaluateurs. En conséquence, ses connaissances doivent ensuite être transmises à cet évaluateur qui s’appuiera dessus pour évaluer le comportement (voir figure 1.3). Pour cela, nous allons aborder les problèmes suivants :

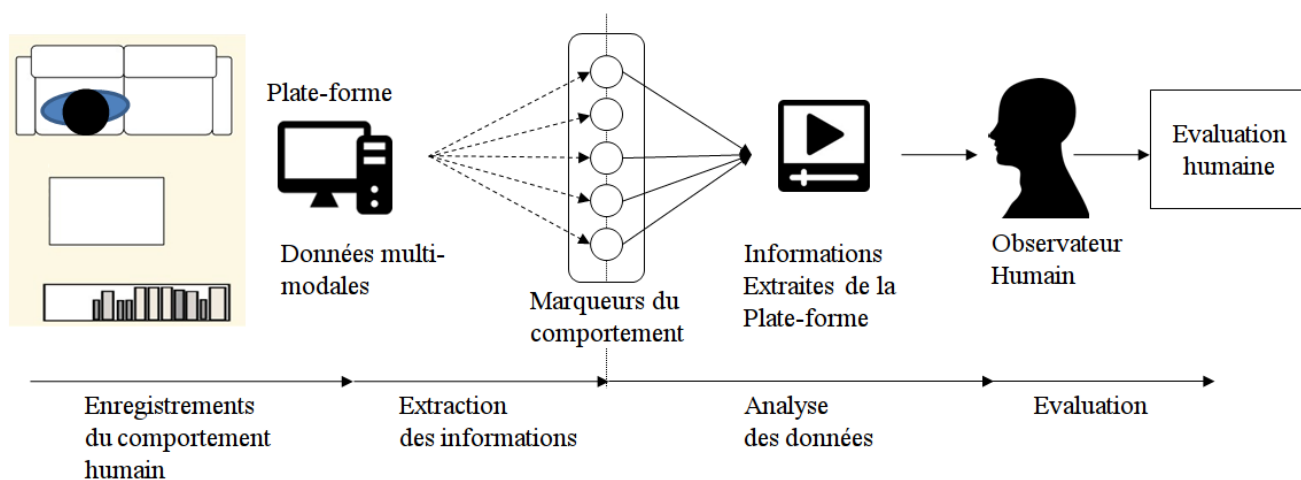


FIGURE 1.3 – Schéma du processus de collaboration entre la plate-forme et un observateur humain

1. L'élaboration d'une plate-forme composée de plusieurs systèmes adaptatifs et modulaires, pour l'observation et l'analyse du comportement humain,
2. La fusion de données multimodales au sein d'un système complexe,
3. La collaboration inter systèmes au sein de la plate-forme pour réaliser une tâche d'observation et d'analyse du comportement humain,
4. L'interaction entre l'humain et la plate-forme pour une évaluation collaborative du comportement,
5. La création d'un corpus d'interactions entre le patient et le soignant lors d'un soin,
6. La méthodologie pour mesurer la performance de la plate-forme et sa collaboration avec un observateur humain.

L'élaboration d'une plate-forme, composée de plusieurs systèmes adaptatifs et modulaires, pour l'observation et l'analyse du comportement humain : Les systèmes mis en place afin de reconnaître les activités sont régulièrement composés d'un système ambiant ou d'un robot mobile [Queirós et al., 2015]. Suite à des limitations matérielles ou à des conditions environnementales variables, des erreurs ou des ambiguïtés dans le traitement des données induisent des résultats faux ou inadéquats. Pour palier à ces problèmes et réaliser la tâche administrée, des systèmes composés de plusieurs robots ou d'un système ambiant et un robot en parallèle peuvent être employés afin de récolter plus d'informations [Panait and Luke, 2005, Hu et al., 2014]. **Nous proposons de mettre en place une plate-forme composée de deux systèmes indépendants, un système ambiant et un robot mobile, capables de collaborer entre eux.** Chaque système fonctionne indépendamment afin de récolter des informations sur l'environnement et le comportement humain. Lorsqu'une information est manquante dans un des systèmes (capteur manquant ou dysfonctionnel, angle mort pour l'observation, etc.), le second système collabore avec le premier pour récupérer l'information. La mobilité du robot complétant l'ubiquité du système ambiant et vice versa.

La fusion de données multimodales au sein d'un système complexe : Cette architecture de plate-forme s'appuie sur des ontologies. Ces dernières, lorsque modélisées à l'aide de fonctionnalités de intrinsèques comme les propriétés d'objets, permettent la fusion de données de haut niveau. Nous nous appuyons sur cette technologie pour standardiser les informations extraites des capteurs. De plus, ces informations extraites de l'environnement et celles déduites des ontologies sont de nature sémantiques et donc lisibles par un être humain. Ceci facilitant l'interaction entre la plate-forme et l'observateur humain pour une évaluation collaborative du comportement.

La méthodologie d'évaluation des performances du système global en collaboration avec un observateur humain est aussi importante pour garantir un niveau de robustesse à l'évaluation de comportement. Les méthodologies d'évaluations du comportement existantes devront être adaptées pour cette plate-forme.

Enfin, la collaboration inter systèmes au sein de la plate-forme pour réaliser une tâche d'observation et d'analyse du comportement humain répond à une problématique de partage de connaissances entre les deux systèmes. Les systèmes sont indépendants mais doivent être en mesure de partager leurs connaissances en temps réel et, au besoin, de récupérer les informations manquantes.

Enfin, pour pouvoir tester, améliorer et valider les algorithmes proposés dans cette thèse nous avons besoin d'un corpus d'enregistrements d'activités et d'interactions entre humains spécifiques aux applications de cette thèse. Au début de notre travail un tel corpus n'existait pas, nous avons donc mis en place des protocoles afin de les réaliser. La méthodologie de l'enregistrement de tels corpus est une autre problématique. Pour effectuer des enregistrements d'une interaction lors d'un soin, l'approbation d'un comité d'éthique et de la

Commission Nationale de l'Informatique et des Libertés (CNIL) doivent être obtenues en amont.

1.4.1 Considérations éthiques pour les expériences et les données numériques

L'étude des comportements humains est un sujet qui requiert de tenir compte de considérations éthiques sur les données recueillies et la méthodologie mise en place pour les recueillir. Tout au long des expériences réalisées dans le cadre de cette thèse, des participants aux études ont été filmés et enregistrés.

Outre les points éthiques engagés par la loi, trois sujets ont représenté le fil rouge de cette thèse :

1. Définir et mettre en place d'une plate-forme éthique dans le cadre de cette thèse et réutilisable pour d'autres travaux,
2. Informer les participants des utilisations des données avant et après les traitements,
3. Engager une procédure de sécurisation des données limitant l'accès des enregistrements,

Nous les détaillons individuellement ci-dessous ainsi que les mesures prises pour y répondre.

Définir et mettre en place d'une plate-forme éthique dans le cadre de cette thèse et réutilisable pour d'autres travaux

Depuis le début des travaux et, à l'instar de la majorité des projets présentés dans la sous section 1.3.1, nous avons décidé que la plate-forme que nous allions mettre en place devait être capable de continuer à fonctionner même si nous décidions de changer ou retirer un type de capteur.

Ce choix de conception a été motivé en premier lieu pour proposer la possibilité à un utilisateur de refuser l'utilisation d'un certain type de capteur. Ainsi nous pouvons le remplacer soit par une autre famille de capteurs, soit simplement le retirer. Cela peut impliquer que la plate-forme fonctionne dans un mode dégradé, mais qu'elle continue de fonctionner malgré tout.

Informers les participants des utilisations des données avant et après les traitements

Les participants à l'étude sont filmés et enregistrés. Ils ont donc été considérés avec respect : respect de leur personne, de leur parole, de leur image et de leur intimité. Pour cela, en amont des enregistrements, ils étaient informés de façon complète et loyale, en des termes compréhensibles, des objectifs et des contraintes de l'étude, des risques éventuels encourus, des mesures de surveillance et de sécurité nécessaires, de leurs droits de refuser de participer à l'étude ou de la possibilité de se rétracter à tout moment.

Ces informations figuraient sur une lettre d'informations distribuée à chaque participant remplissant les critères d'inclusion (voir annexes E, F). Pour chaque participant, le consentement libre et éclairé a été recueilli avant l'inclusion définitive dans l'étude. Le protocole et les lettres d'informations ont été soumis pour avis à un comité d'éthique et un Comité de Protection des Personnes.

Suite aux expériences, des résultats ont été obtenus. Nous les avons communiqué aux soignants qui nous ont accompagné lors des expériences ainsi qu'à leur collègues afin de les

sensibiliser aux questions de recherches. Nous avons pris le temps de répondre à leurs questions et leur crainte de voir la robotique remplacer l'humain dans le domaine médical. Cette thèse n'a pas pour objectif de remplacer l'être humain dans une tâche mais d'optimiser sa réalisation en transférant une partie de l'évaluation a posteriori et en anonymisant les informations des participants. Ainsi, nous souhaitons proposer une nouvelle approche pour réaliser des évaluations de comportement en collaboration avec un humain plutôt que le remplacer.

Engager une procédure de sécurisation des données limitant l'accès des enregistrements

Concernant les données numériques des enregistrements des interactions entre différentes personnes, nous avons répondu à une étude de risques demandée par la CNIL où nous avons proposé des mesures pour sécuriser les données :

1. Les données sont enregistrées dans un disque dur crypté et protégé par un mot de passe. Ainsi, si le disque dur était amené à être volé, l'accès aux données est complexifié au maximum,
2. Les enregistrements sont réalisés dans un hôpital, le traitement au sein des laboratoires de l'ISIR. Le point précédent assure une première sécurité en cas de perte des données lors du transfert, en revanche, au sein de l'hôpital ou de l'ISIR, le disque dur est stocké dans une salle fermée à clé et uniquement accessible par des personnes habilités,
3. Le disque dur est consulté sur un ordinateur ayant un OS sous noyau linux et n'ayant pas directement accès à internet. De plus, les données restent en local sur le disque dur chiffré à l'aide d'un algorithme de chiffrement symétrique (AES128) et protégé lui aussi par un mot de passe,
4. Les données initiales (avant traitement) sont sauvegardées sur un disque dur externe, comme précédemment, protégé par mot de passe et chiffré à l'aide d'un algorithme de chiffrement AES128.
5. Enfin, l'accès aux données est restreint aux seules personnes autorisées par le promoteur de l'étude.

L'ensemble de ces mesures permet une protection efficace et maximale contre la perte ou le vol de données. Cette procédure complète a été validée par la CNIL.

Conclusions

Nous avons mis en place une plate-forme modulable comprenant un ensemble de capteurs. L'emploi de certains types de capteurs peut gêner ou rebuter des utilisateurs. Ainsi, dès le début de sa conception, la caractéristique permettant l'ajout ou le retrait de certains d'entre eux a été établie au sein de la plate-forme.

Les utilisateurs et participants des différentes études réalisées tout au long de ces travaux de thèses ont été, autant que possible, informés des tenants et aboutissants de l'expérience. De même lorsque les résultats ont été obtenus, nous avons cherché à les communiquer afin de populariser nos travaux de recherche, nos intentions et ce que nous découvrons à travers les expériences pratiquées avec leur participation ou celle de leur collègues.

Finalement, ces études ont permis d'obtenir des données qui restent sensibles. Nous avons donc mis en place une procédure permettant de les sécuriser et les rendre, autant que possible, inaccessible à une personne non autorisée.

1.5 Organisation du document

Au vu des objectifs et des problématiques de la section précédente, nous proposons l'organisation suivante pour ce manuscrit de thèse. La plate-forme est détaillée dans le chapitre 2. Les systèmes, algorithmes et capteurs employés y sont décrits ainsi que la méthodologie de collaboration entre le système ambiant et le robot mobile pour le partage d'informations environnementales manquantes.

Ce document continue sur un chapitre 3 consacré à l'analyse automatique des activités de la vie quotidienne par la plate-forme et les observateurs humains. Nous commençons par présenter l'étude que nous avons menée et l'adaptation de la plate-forme pour les observations des activités. Enfin, nous analysons les performances des observateurs extérieurs s'appuyant sur les informations de la plate-forme. Deux méthodes sont alors employées : la première ne permettant pas de mettre en évidence la capacité des observateurs à observer les comportements, une seconde a été exploitée. Cette dernière permet de visualiser les performances réelles des observateurs à reconnaître un comportement en utilisant uniquement les informations de la plate-forme.

Le chapitre 4 traite de l'évaluation automatique de l'interaction entre un patient et un soignant au cours d'un soin. Après une brève introduction, nous détaillons la méthodologie mise en place ainsi que l'adaptation employée de la plate-forme décrite dans le chapitre 2. Ce chapitre est décomposé en deux expériences, l'une réalisée en conditions contrôlées et la seconde en conditions réelles. Pour chaque étude, nous comparons des évaluations d'observateurs humains extérieurs aux évaluations faites par la plate-forme en collaboration avec un humain afin de mesurer leur performance.

Le manuscrit se termine par les conclusions et les perspectives envisagées suite à ces travaux dans le chapitre 5.

Chapitre 2

Une plate-forme de modélisation du comportement

Nous présentons dans ce chapitre l'architecture d'une plate-forme de capteurs multimodales. Cette plate-forme dispose de plusieurs caractéristiques comme la modularité et l'interopérabilité des sous systèmes qui la compose. Tout au long des travaux réalisés dans le cadre de cette thèse, nous utilisons soit directement la plate-forme à son potentiel complet, soit une variante plus optimisée pour les besoins de l'étude. Dans les prochaines sections, outre l'aspect matériel et algorithmique, nous présentons les sous systèmes et la méthodologie implémentée afin qu'ils collaborent entre eux pour réaliser des tâches qui leurs sont assignées.

2.1 Description de la plate-forme

2.1.1 Un système de systèmes

La plate-forme permettant l'observation et l'évaluation de comportement est composée de deux systèmes. Le premier est un système ambiant composé de capteurs récoltant des informations environnementales et comportementales (détection de mouvement, position d'humains, capteurs reconnaissant les étapes du sommeil, etc.). Le second est un robot mobile comportant uniquement une caméra pour récolter un ensemble de données plus restreint que le système ambiant concernant l'environnement et le comportement (détection de mouvements, position et reconnaissance d'humains, etc.). Chaque système possède ses propres capteurs et traite uniquement les informations que ces derniers extraient (voir schéma traitement de l'information 1.2). Ces deux systèmes sont donc entièrement autonomes et peuvent fonctionner sans que l'autre ne soit activé.

Le système ambiant a pour fonction principale de récupérer des informations de l'environnement en permanence. Le robot, quant à lui, a deux fonctionnalités. La première lui fait récupérer des informations de l'environnement en permanence; la seconde implique la recherche d'informations précises et manquantes du système ambiant. Ces deux modes de fonctionnement ne rentrent pas en conflit entre eux et peuvent être employés en même temps.

Le système ambiant est modulable afin de pouvoir intégrer les capteurs nécessaires pour l'observation du comportement ciblé. Par exemple, dans le cas d'analyse des activités de la vie quotidienne, un capteur de sommeil, des capteurs de mouvement et des caméras sont nécessaires (voir la sous section 3.1.3 pour plus d'informations); alors que dans le cas d'analyses de l'interaction entre un patient et un soignant lors d'un soin, seuls des caméras

et microphones sont nécessaires (voir la section 4.1.3).

Afin de faciliter la mise en place de différents types de capteurs au sein du système ambiant, nous avons besoin d'utiliser un ensemble d'outils informatiques facilitant le développement de logiciels pour la robotique intégrant dans ses fonctionnalités de base les principes de :

- cross-plate-forme¹,
- communication inter-processus,
- modularité de chaque fonctionnalité développée,
- intégration d'un grand nombre de capteurs.

Ainsi, des drivers logiciels pour chaque capteur ont été développés. Ces drivers traitent les signaux reçus des capteurs (comme les images, sons, etc.) et extraient des données (identification de la personne, lieux du mouvement, etc.). Ajouter un capteur au système nécessite donc de lui créer, ou réutiliser un driver adéquat.

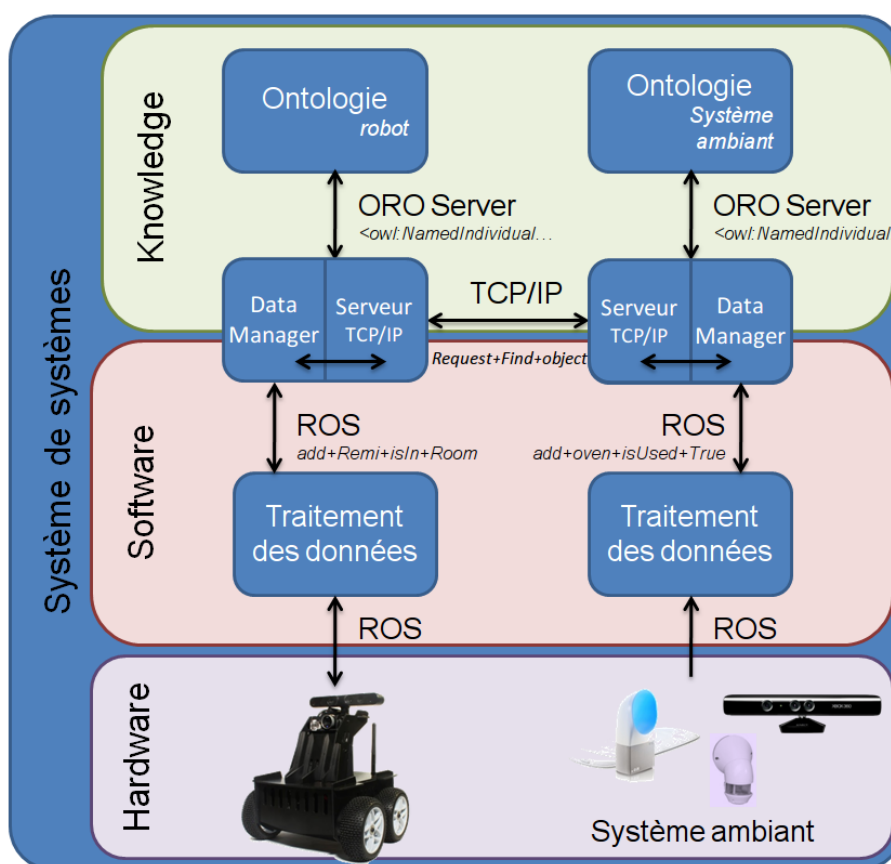


FIGURE 2.1 – Architecture électronique et informatique de la plate-forme

La plate-forme ainsi composée d'un système ambiant et d'un robot s'articule autour d'une architecture détaillée dans le schéma 2.1. Pour chaque système la chaîne d'analyse des informations est la même : Les capteurs récupèrent des données qui sont transmises à un ordinateur. Des algorithmes sont ensuite appliqués pour extraire des informations comme le positionnement d'un humain dans l'environnement, son activité en cours, etc. Ces informations sont transmises et stockées au sein d'une base de données.

Peu importe le capteur ou le système d'où ces informations proviennent, elles doivent être formalisées de manière à pouvoir être comparées entre elles (voir la sous section 1.2.1). De plus, nous souhaitons que les données stockées soient immédiatement intelligibles par

¹ possibilité d'utiliser tout ce qui a été déjà fait sur une plate-forme directement sur une plate-forme

un utilisateur de la plate-forme. Ainsi, de par leurs propriétés sémantiques et logiques, et leur utilisation dans les domaines de la reconnaissance d'activité [Riboni and Bettini, 2011, Chen et al., 2012], nous avons fait le choix d'utiliser des ontologies comme base de données.

2.1.2 Logiciels et algorithmes employés

Les capteurs renvoient des signaux à l'ordinateur qui sont ensuite traités par des algorithmes chargés d'extraire des informations. La contribution de ce travail porte sur la mise en place de la plate-forme collaborative entre l'humain, le robot et le système ambiant. Les algorithmes employés proviennent de la littérature :

Fonctionnalité de l'algorithme	Référence
Extraction du squelette depuis la kinect	OpenNI Kinect Library
Localisation de sons	HARK sound localization
Détection des phases du sommeil	Aura, Withings© [Nakadai et al., 2008]
Détection d'éléments dans une image (humains, sourires)	Viola and Jones [Viola and Jones, 2001]
Mesure du niveau de stress	Quantité de Mouvement [Aigrain et al., 2016]

TABLEAU 2.1 – Tableau des algorithmes employés et leur référence

La thèse ayant pour objectif de mettre en place une plate-forme d'aide à l'évaluation du comportement, les algorithmes employés proviennent de la littérature et ne sont pas évalués ici. Ces algorithmes ont été choisis pour leur représentation et leur utilisation courante dans la littérature.

Une fois les informations extraites, elles sont transmises à un gestionnaire de données employé pour mettre à jour l'ontologie du système correspondant ou les partager avec l'autre système.

Ici, l'ontologie se comporte comme une base de données sémantique et logique. En appliquant des règles et des propriétés définies en amont des informations supplémentaires peuvent être extraites et représenter l'environnement perçu par le système (voir la sous section 1.2.2).

2.2 Processus de collaboration inter-système

2.2.1 Description du processus

Précédemment dans ce chapitre, le rôle et le fonctionnement de chaque système a été présenté. Durant l'accomplissement d'une tâche par un système, des informations peuvent venir à manquer. A ce moment, les deux systèmes intégrés dans la plate-forme collaborent entre eux afin de réaliser la tâche attribué.

Par exemple, lorsque le système ambiant détecte un mouvement dans la cuisine, sans caméra placé à cet endroit, il ne peut déterminer qui est présent et quels sont les outils qu'il emploie (four, réfrigérateur, etc.). Le système ambiant, demande alors au robot s'il possède des données concernant la personne présente dans la cuisine.

Le processus de collaboration est initié par le gestionnaire de données du système où la donnée est manquante. En termes informatiques, le système qui a détecté l'erreur effectue une requête auprès du second système. Chaque requête est stockée dans une file (First In First Out) de manière à ce qu'elle soit traitée après que la précédente soit elle-même traitée. Quand un système fait une requête à l'autre, la donnée en question est recherchée dans l'ontologie du deuxième système. Si elle est trouvée, elle est renvoyée au système qui a émis la requête.

Dans le cas d'une demande faite au système ambiant par le robot. Si celui-ci n'a aucune donnée à renvoyer au robot, il va donc simplement renvoyer que "rien n'a été trouvé". À l'inverse, dans le cas d'une demande envoyée au robot par le système ambiant, si celui-ci n'a aucune donnée à renvoyer, il va physiquement se déplacer à la recherche de l'information manquante. En effet, chaque donnée dépend directement d'un lieu : lieux de l'activité, emplacement d'un objet, d'une personne, etc. Pour chaque lieu il existe un repère sur la carte de l'environnement du robot. C'est de cette manière que le robot sait où chercher l'information demandée par le système ambiant. En cas d'impossibilité pour le robot à récupérer cette information même après son déplacement, il renverra au système ambiant que "rien n'a été trouvé", sinon il renverra la nouvelle information fraîchement récoltée.

2.2.2 Aperçu du processus à travers un scénario

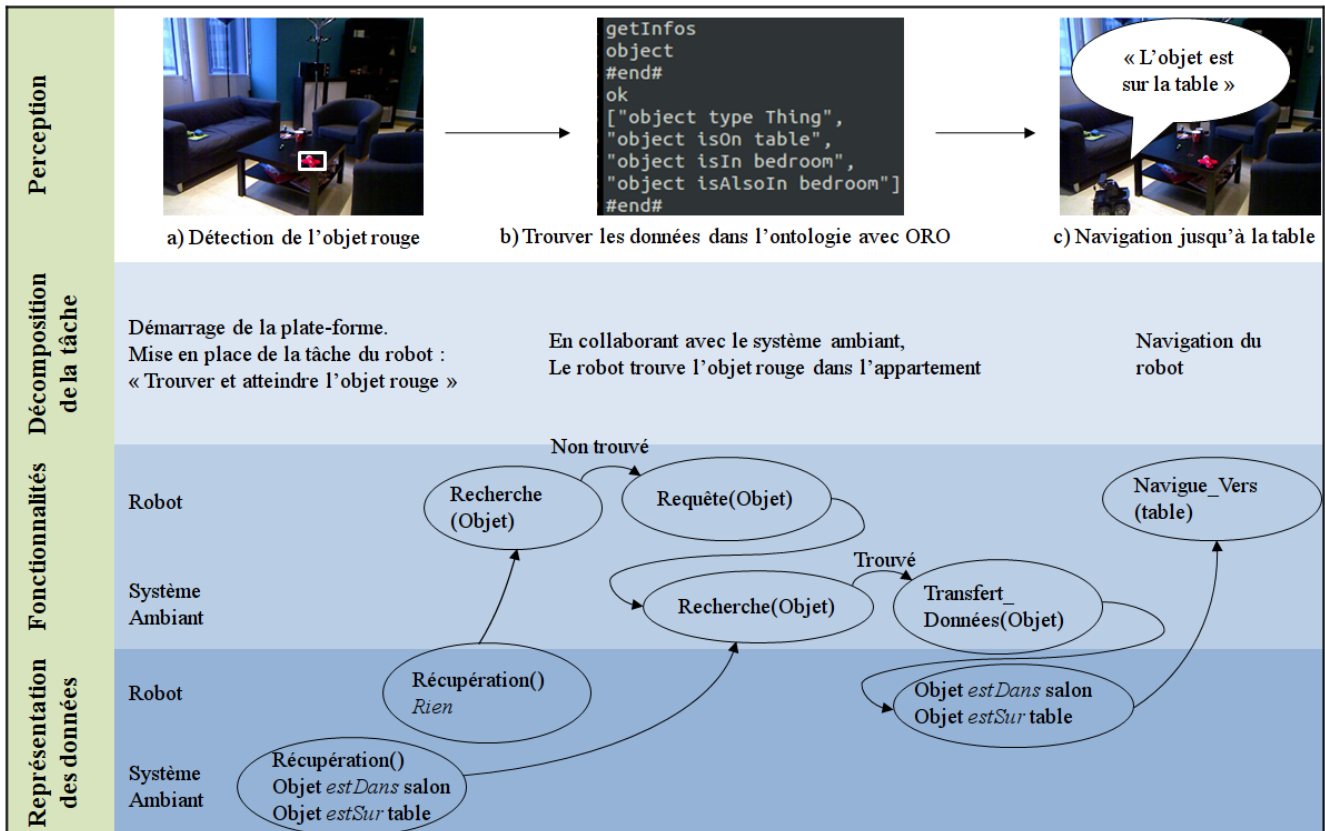


FIGURE 2.2 – Schéma du processus de collaboration inter-système pour de la recherche d'objet

Pour illustrer le processus de collaboration au sein de la plate-forme, un scénario de recherche d'un objet, où la collaboration entre le robot et le système ambiant est nécessaire, a été mis en place. Dans ce scénario, la coopération entre le système ambiant et

le robot est mise en œuvre suite à la recherche d'un objet par un utilisateur humain. Ce dernier demande au robot de la plate-forme de trouver un objet rouge spécifique et de l'y emmener. Dans ce scénario, aucun système n'a volontairement les mêmes capteurs.

Deux précautions sont prises dans ce scénario forçant la collaboration entre les deux systèmes : La détection de couleur n'est possible qu'en employant le système ambiant et aucune information concernant cet objet de couleur rouge n'est présente dans l'ontologie du robot. De plus, cette fonctionnalité est la seule implémentée au sein du système ambiant qui cherchera automatiquement à localiser en permanence cet objet rouge. Les interactions entre les deux systèmes sont entièrement transparentes pour l'utilisateur qui n'a aucune information sur d'éventuels échanges entre les deux systèmes.

Le schéma 2.2 récapitule les interactions entre les systèmes suite à la demande de l'utilisateur :

1. La première étape qui y est décrite consiste à lancer la plate-forme et définir l'objectif ; ici "Trouver et atteindre l'objet rouge". L'objet en question est posé sur une table dans un environnement contrôlé. Il est le seul élément rouge visible par les caméras du système ambiant volontairement ; le principe étant ici de constater la collaboration entre les deux systèmes.
2. Dans un deuxième temps et après les initialisations, le système ambiant extrait automatiquement les informations de l'environnement et détecte l'objet rouge. Il enregistre des données sous forme de triplets "*objet est rouge*" ; "*objet estSur table*" ; "*objet estDans salon*" directement au sein de son ontologie.
3. De son côté, le robot récupère toutes les informations contenues au sein de son ontologie concernant l'objet rouge. Comme l'ontologie ne possède aucune information à ce sujet, il envoie une requête de clarification auprès du système ambiant concernant l'objet rouge.
4. Du point de vue des caméras du système ambiant, donner les coordonnées absolues de cet objet au robot n'est pas envisageable puisque la perspective du robot est différente. En conséquence, les données cartésiennes de cet objet ne sont pas requises tandis qu'à l'inverse son positionnement relatif dans l'environnement permet de donner une information utile au robot. Ainsi, lors de la réception de la requête de clarification, le système ambiant va renvoyer en réponse au robot l'intégralité des informations concernant l'objet rouge demandées : "*objet est rouge*" ; "*objet estSur table*" ; "*objet estDans salon*". Ces informations reçues par le robot vont être directement ajoutées à sa propre ontologie.
5. A ce moment, les ontologies du robot et du système ambiant possèdent les mêmes informations sur l'objet. Le robot est donc en mesure de réaliser l'objectif de trouver et de se rendre auprès de l'objet. Il va donc se déplacer au plus près de la table et interagir avec l'utilisateur en lui spécifiant que "*objet estSur table*".

2.2.3 Spécifications spatiales au sein de l'ontologie

La collaboration entre le système ambiant et le robot, des informations de localisation sont échangées. Ce scénario illustre les ambiguïtés spatiales qui peuvent apparaître quand les points de vue des deux systèmes ne sont pas partagés. Ainsi, dans notre approche, des objets considérés comme n'ayant pas à bouger de manière journalière sont décrits dans l'ontologie lors de son initialisation. Comme des échanges d'informations spatiales doivent avoir lieu, un point de vue global a été mis en place au sein des deux systèmes à l'aide des propriétés suivantes :

[Objet] estDans [Pièce]. Un objet ou un humain est détecté dans une pièce spécifique. Cette propriété est fonctionnelle, ce qui signifie, par exemple, qu'une personne ne peut être dans deux lieux différents.

[Objet] estSur [Objet]. Un objet ou un humain est au dessus d'un autre objet lorsque qu'un capteur détecte le premier à un endroit où un objet connu comme étant stationnaire devrait être, par exemple : une table, un lit, un canapé, etc. Cette propriété est fonctionnelle.

[Objet] estProcheDe [Objet]. Un objet ou un humain est à côté d'un autre humain ou objet quand la distance métrique entre les deux est considérée comme petite. Cette propriété est symétrique.

2.3 Choix techniques

Tout au long de la mise en place de cette plate-forme, des choix techniques ont été réalisés en fonction des contraintes techniques qui s'imposaient.

Afin de faciliter la mise en place de différents types de capteurs au sein du système ambiant, nous utilisons ROS. ROS (Robotic Operating System [Quigley et al., 2009]) est un ensemble d'outils informatiques open source facilitant le développement de logiciels pour la robotique et répondant à nos besoins listés dans la sous section 2.1.1 [web, 2017].

Ainsi, le robot mobile que nous employons est un robot de télésurveillance nommé "*oculus prime*" commercialisé par la société Xaxxon©. La société Xaxxon met à disposition un ensemble de logiciels, fonctionnant sous ROS, permettant une prise en main facilitée du robot. Le robot apporte de la mobilité au système ambiant. Ainsi, une des fonctionnalités primordiales pour le robot est sa capacité à se déplacer dans l'environnement dans lequel il est placé. L'environnement est tout d'abord cartographié et étiqueté en utilisant l'algorithme SLAM (Simultaneous Localization And Mapping) [Durrant-Whyte and Bailey, 2006] à l'aide du capteur infrarouge Orbbec©Astra. Le robot sait donc où se situent des lieux précis comme *le salon, la cuisine, la chambre* mais aussi le positionnement du *réfrigérateur, du canapé*, etc.

Enfin, le robot dispose d'une caméra et d'un haut parleur. La caméra est employée pour détecter un être humain et le haut parleur permet d'exprimer des informations à l'humain (voir la section 2.2).

Le gestionnaire de données employé pour mettre à jour l'ontologie du système correspondant s'appuie sur le serveur de la plate-forme open source "Open Robot Ontology" (ORO) [Lemaignan et al., 2010] :

ORO est une plate-forme open source qui s'appuie sur le langage RDF. Elle inclut, entre autre, les fonctionnalités permettant d'ajouter ou supprimer des éléments, de rechercher des concepts ou encore de vérifier la cohérence de l'ontologie. La mise à jour de l'ontologie se fait par l'exécution de requêtes SPARQL (SPARQL Protocol and RDF Query Language - Langage des Protocoles et Requêtes RDF). L'utilisation d'ORO permet de simplifier l'ajout d'informations au sein de l'ontologie par rapport aux requêtes SPARQL. De plus, ORO utilise le raisonneur Pellet [Sirin et al., 2007] en temps réel. Ainsi, à chaque fois qu'une donnée est ajoutée au sein de l'ontologie, ORO vérifie sa cohérence, ceci afin de détecter immédiatement les erreurs et les éléments qui y sont impliqués. Si une erreur est détectée, le dernier ajout de données est annulé et un message d'erreur est renvoyé.

Le partage des informations, quant à lui, se fait en employant une simple connexion

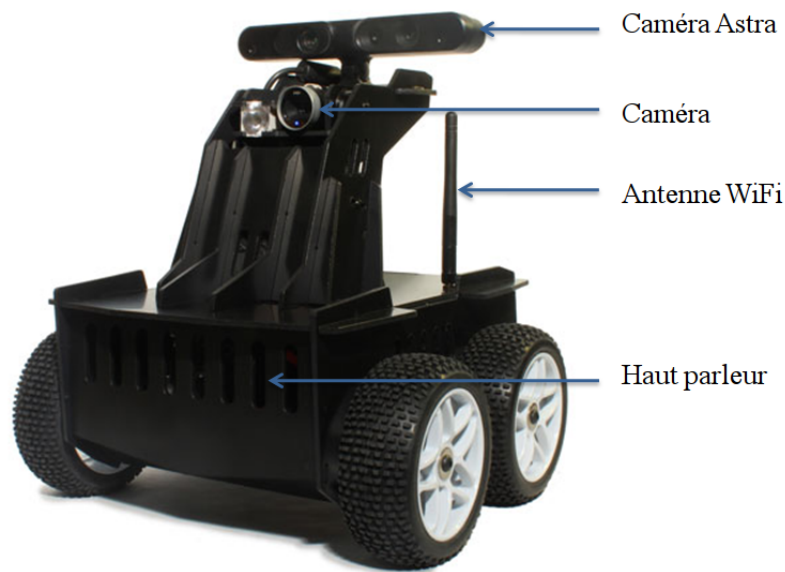


FIGURE 2.3 – Robot mobile oculus prime

TCP/IP² par le biais de serveurs/clients chargés des échanges de requêtes. Quand un serveur en reçoit une, il la transfère au gestionnaire de données en utilisant les outils de communication ROS. Finalement, la demande est traitée et la réponse est retransmise au client via le serveur.

2.4 Conclusion

Dans ce chapitre, une plate-forme composée d'un système ambiant et d'un robot mobile a été présentée. Cette plate-forme a la particularité d'être modulable, c'est à dire qu'elle continuera de fonctionner suite à l'ajout ou le retrait de capteurs. Chaque sous système est en mesure de collaborer avec un autre grâce à l'emploi d'ontologies permettant de formaliser des données extraites d'informations de capteurs. Cette formalisation permet l'interopérabilité de chaque données avec n'importe quelle autre sans se soucier de sa provenance. Enfin, nous avons mis en place un processus de collaboration entre les systèmes. Ce processus se met en place lorsque des données sont manquantes chez un système, il peut alors les demander auprès d'un autre afin de réaliser les tâches qui lui sont attribuées.

² N'importe quel protocole de réseau plus sécurisé aurait pu être utilisé à la place, nous avons choisi celui-ci pour sa simplicité et rapidité à être déployé à travers un réseau Wi-Fi sur lequel seul le robot et le système ambiant sont connectés.

Chapitre 3

Collaboration entre un robot et un système ambiant pour la reconnaissance d'activités

Dans ce chapitre, nous souhaitons démontrer que la collaboration entre un observateur humain et une plate-forme automatique de reconnaissance d'activités de la vie quotidienne permet d'observer les comportements de manière équivalente à un observateur présent lors de ces activités.

Pour cela, nous avons mis en place une expérience où la plate-forme automatique est chargée d'extraire des informations de l'environnement pour reconnaître une partie des activités. Dans un second temps, un observateur complète la reconnaissance d'activités à partir des informations extraites par le système. Cet observateur ne s'appuie que sur des données sémantiques (voir la section 2.1) afin de terminer la reconnaissance. Enfin, nous comparons les observations ainsi obtenues à celles faites par l'observateur présent.

3.1 Étude sur la reconnaissance des activités de la vie quotidienne

3.1.1 Objectif et hypothèse de l'étude

L'objectif de cette étude est de mesurer la capacité qu'a un observateur à reconnaître les activités de la vie quotidienne d'un humain à partir des enregistrements et informations délivrées par la plate-forme. Cette étude se décompose en plusieurs étapes :

1. L'enregistrement par la plate-forme in situ des activités de la vie courante d'un humain,
2. La reconnaissance des AVQ par deux observateurs employant les enregistrements réalisés précédemment. Ces observations établissent la vérité terrain,
3. La reconnaissance des AVQ par des observateurs employant les informations extraites par la plate-forme automatique depuis les enregistrements réalisés précédemment,
4. La comparaison des observations réalisées lors des deux étapes précédentes.

Durant cette étude, deux types d'évaluations sont réalisées :

La première, est faite par plusieurs Observateurs Référénts (OR) qui utilisent une vidéo avec une vue d'ensemble sur toutes les activités. Ils sont chargés d'annoter le séquençement et les horaires des différences activités. Leur évaluation est employée comme référence et sert de point de comparaison avec les résultats obtenus lors de la seconde évaluation. La seconde évaluation est réalisée par des individus, sans critères d'inclusions, participant à l'étude : les Observateurs Participants (OP). Seules des informations sémantiques provenant des ontologies de la plate-forme leurs sont proposées pour reconnaître les activités.

L'hypothèse posée est que la plate-forme permet un recueil d'informations équivalent aux informations que peut avoir un observateur ayant accès à l'ensemble des enregistrements vidéos et sonores de l'environnement dans lequel les activités sont faites. Les résultats attendus sont que l'apport de la technologie permet de fournir des moyens de perception à distance fiables pour l'analyse des activités permettant ainsi d'éviter soit la présence d'un observateur durant les activités de la vie courante, soit la visualisation de l'environnement par l'utilisation d'enregistrements vidéo.

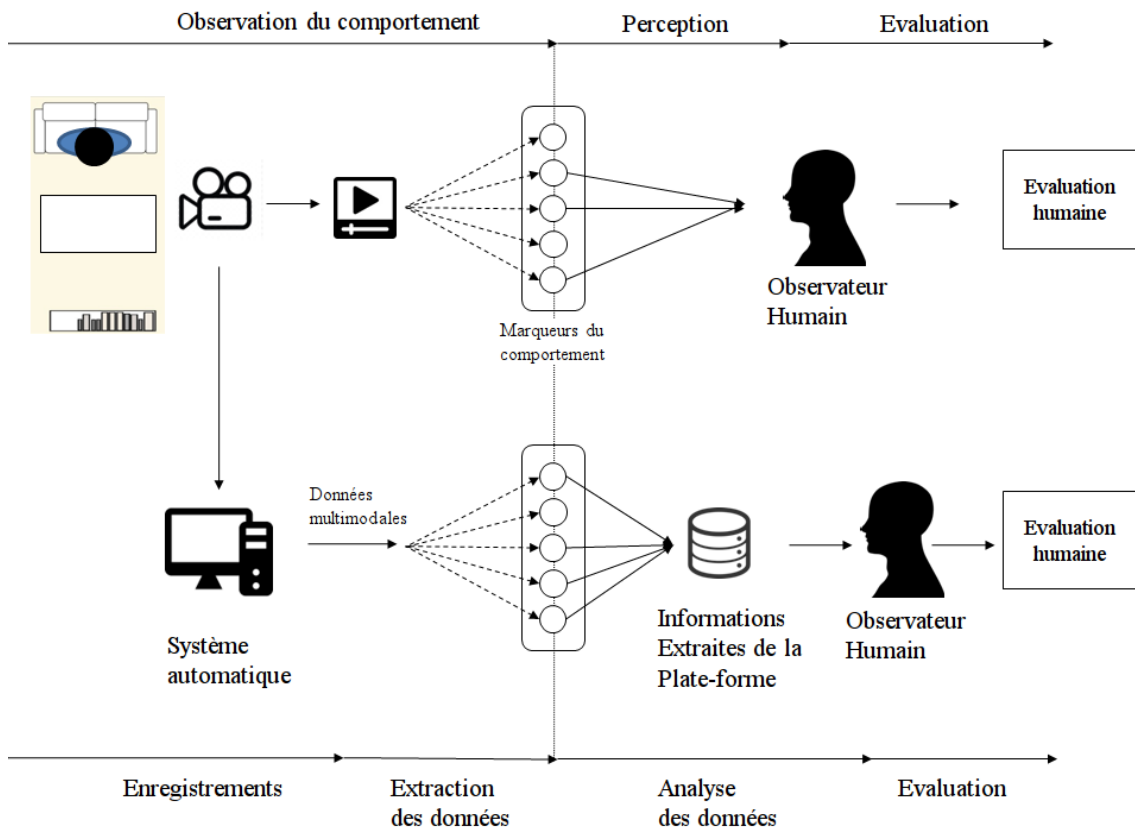


FIGURE 3.1 – Schéma de l'étude. L'Observateur Référént (OR) annote les activités en cours. L'observateur Participant (OP) utilise les informations extraites par la plate-forme afin de reconnaître et annoter les activités observées.

3.1.2 Schématisation de l'observation des activités

Le sujet humain réalisant les activités est placé dans un environnement proche d'un réel appartement. Cet environnement comporte trois pièces : une chambre, une cuisine et un salon. Le sujet a reçu en amont une fiche d'informations comprenant la liste des activités à réaliser et des informations sur comment les réaliser :

- Cuisiner
- Dormir
- s’Habiller
- Manger
- Regarder la télévision
- Téléphoner

Le sujet a donc un set de plusieurs activités à réaliser de manière séquentielle (une activité après l’autre) ou simultanée (deux activités ou plus en parallèle).

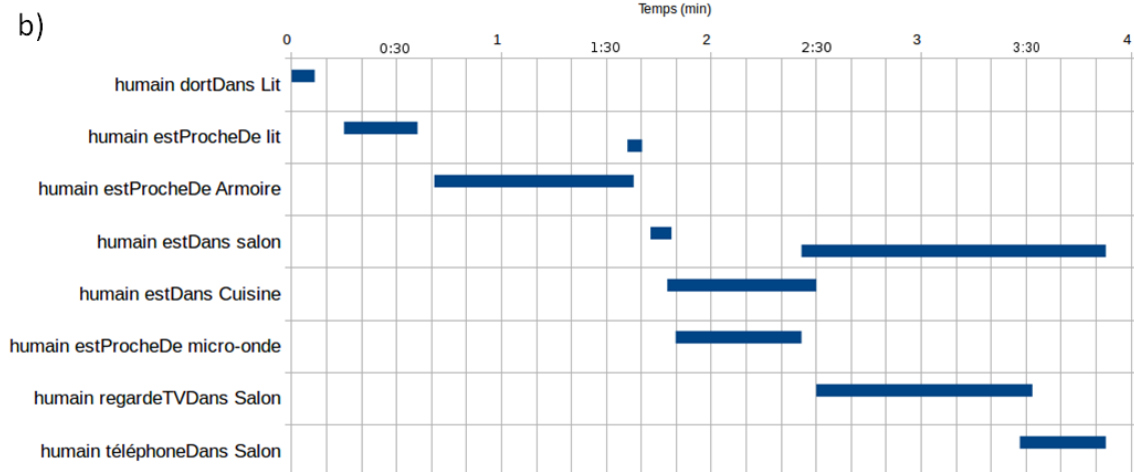
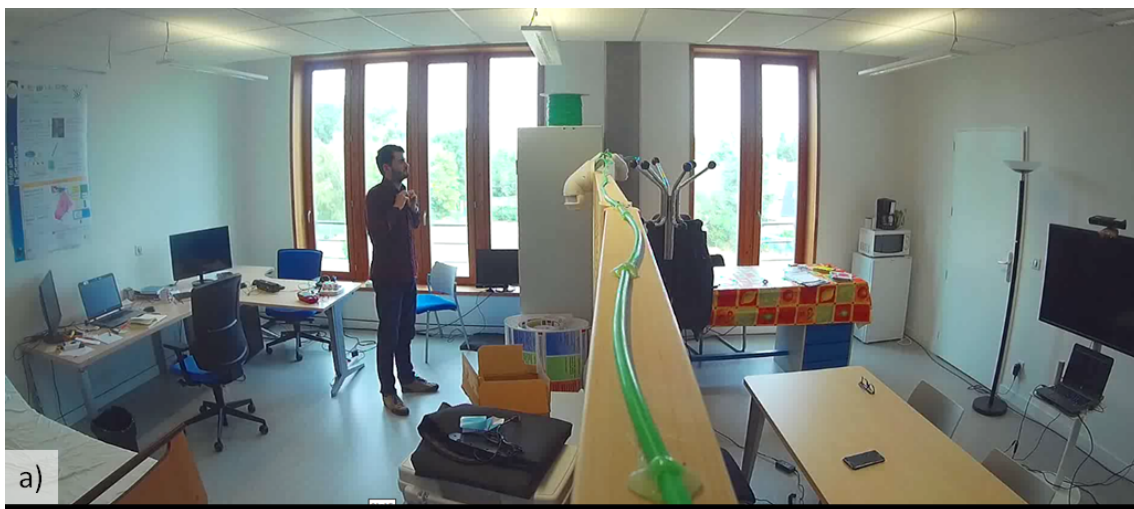


FIGURE 3.2 – a) Aperçu de la vue globale des OR ; b) Aperçu des informations que les OP ont pour évaluer les activités

L’enregistrement dure environ 5 minutes selon le temps nécessaire au sujet pour réaliser les activités. Tout au long de la scène, la plate-forme est activée et les signaux reçus par chaque capteur sont enregistrés, pour permettre l’analyse automatique à posteriori. La figure 3.2 montre le point de vue qu’ont les observateurs référents ainsi que toutes les informations délivrées aux observateurs participants pour annoter les activités de la vie quotidienne.

Durant les observations des observateurs référents, ces derniers annotent un tableau avec les différentes activités, l’heure de début et l’heure de fin correspondante. Les observateurs participants ont accès uniquement à des informations récoltées par les capteurs. Il leur est

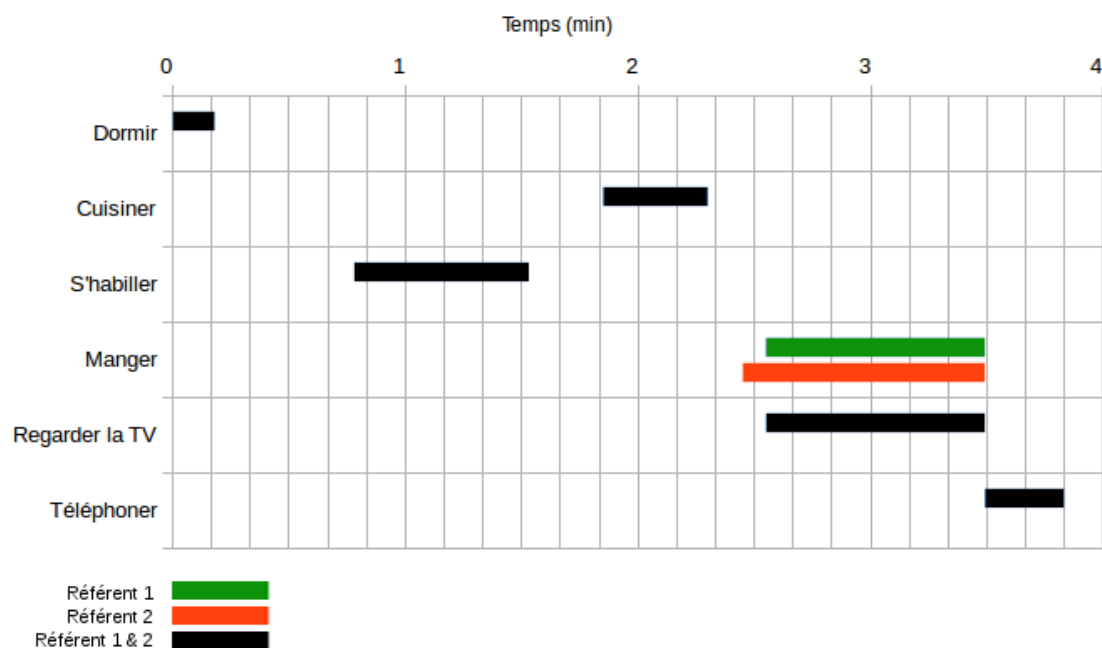


FIGURE 3.3 – Séquencement des horaires des activités observées par les observateurs référents

demandé pour chaque activité de renseigner son heure de début, son heure de fin et sur quelles informations provenant des capteurs ils se sont appuyés pour donner leurs réponses. Hormis pour l'activité s'habiller où il y a un décalage de 10 secondes entre les deux horaires de début, les OR ont donné à deux secondes près les mêmes temps (voir la figure 3.3). En conséquence, et afin de simplifier la tâche consistant à dater le début et la fin d'une activité pour les OP, des intervalles de 10 secondes ont été utilisés pour les OP au moment de renseigner ces horaires. En outre, cet intervalle de 10 secondes permet aussi de simplifier les comparaisons entre les horaires des OR et des OP.

3.1.3 Caractéristiques techniques et algorithmique

Le système ambiant employé pour cette étude a donc été adapté en fonction des activités que nous cherchons à observer. Ainsi, il est composé d'un capteur de sommeil, de capteurs infrarouge pour la détection de mouvement et d'une caméra kinect pour mesurer l'attention que porte le participant à la télévision (voir la figure 3.4 pour plus de détails). De plus, sans être réellement un capteur, une application Android placée sur le téléphone portable afin de détecter le début et la fin d'un appel téléphonique. Enfin, le robot mobile embarque avec lui une caméra.

Durant l'analyse des AVQ chaque système collabore avec l'autre afin de compléter les informations manquantes.

- Les capteurs de mouvements infrarouge (legrand©), *utilisés comme détecteur de mouvement, ils indiquent le mouvement d'une personne dans une pièce* ;
- Les caméras Kinect (Microsoft©), *utilisées ici comme caméra RGB pour la localisation d'humains dans l'espace* ;
- Le capteur de sommeil Aura (Withings©), *capteur placés au niveau du lit qui détecte et enregistre les différentes phases de sommeil*.

Pour la détection d'humain dans l'environnement à l'aide des caméras, nous nous appuyons sur l'algorithme proposé par Viola and Jones [Viola and Jones, 2001]. Aura est un dispositif développé par la société Withings©, nous n'avons pas accès aux informa-



FIGURE 3.4 – Capteurs employés dans le système ambiant : a) Kinect, b) Capteur de Sommeil, c) Capteur de mouvement infrarouge

tions concernant les algorithmes employés pour la détection et reconnaissance de cycle du sommeil. En revanche, ce capteur connecté possède une API permettant de récupérer les informations utiles du capteurs. Enfin les capteurs infrarouge détectent uniquement un mouvement lorsque celui-ci se produit. Ils renvoient alors une simple information indiquant qu'ils ont détectés un mouvement. En fonction de leur emplacement, nous savons quelle zone ils observent et donc dans quelle zone a eu lieu le mouvement. De son côté, le robot, à l'aide de sa caméra embarquée, détecte les mouvements dans la cuisine. Dans notre cas, ces mouvements seront ceux réalisées par le participant se déplaçant vers le micro-onde. Lorsque des mouvements sont détectés, le robot suit la personne et détermine, grâce à son positionnement dans l'espace, à côté de quel objet de cuisine la personne s'est arrêtée (table ou micro-onde).

Dans la figure 3.6, les capteurs sont mis en relation avec les informations qu'ils extraient de l'environnement. Dans un premier, seul trois capteurs (Aura, le téléphone et la kinect) indiquent une information directement en lien avec une des activités pouvant être réalisées. Les informations de positionnement dans l'espace sont données par les capteurs infrarouge ou le robot. Ce dernier ne s'affairant que dans la cuisine, les informations qu'il extrait ne concerne que ce lieu.

Propriété d'objets	Sous propriétés chaînées
téléphoneDans	téléphoneAvec [Téléphone] o [Téléphone] estDans
regardeTvDans	regarde [TV] o [TV] estDans

TABLEAU 3.1 – Tableau des propriétés chaînées utilisées au sein de l'ontologie

Deux propriétés d'objets sont employées pour décrire plus précisément les activités en spécifiant leur lieux de réalisation *téléphoneDans* et *regardeTVDans* (voir tableau 3.1). Ces deux propriétés sont déduites via une propriété chaînée, respectivement (*téléphoneAvec o estDans*) et (*regarde o estDans*). Outre une utilité limitée dans notre scénario, ces chaînes de propriétés démontrent l'utilisation possible de raisonnement plus détaillé utilisable au sein de l'ontologie.

Aura est un capteur permettant de donner des informations sur le sommeil de son utilisateur via une API. Les informations récupérables sont donc directement l'horaire de début et l'horaire de fin du sommeil de l'utilisateur.

Les capteurs infrarouge ont une zone de détection (en bleu sur la figure 3.5). Lorsqu'un mouvement est détecté dans la zone correspondante une information de positionnement est relayée au sein de l'ontologie. Ainsi, le capteur dans la chambre indique qu'un *humain estACotéDe armoire*, le capteur à l'entrée du salon indique *humain estDans salon* et enfin,

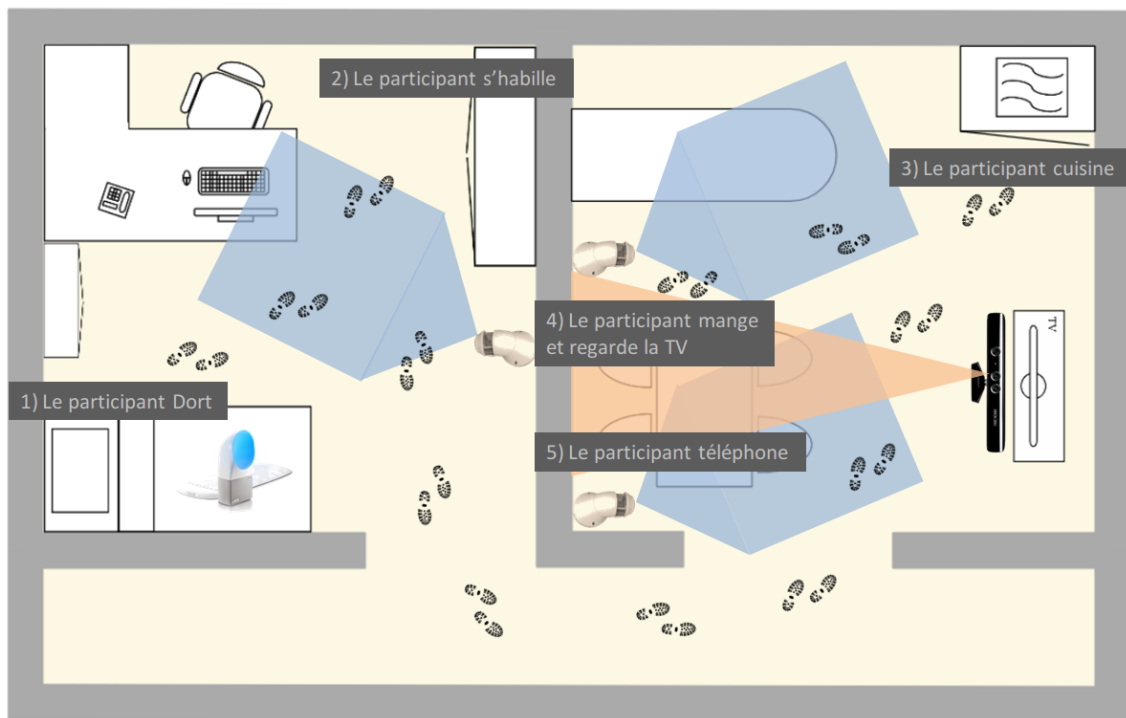


FIGURE 3.5 – Disposition du système ambiant dans l’environnement et séquence-
 ment des activités. 1) *Le participant dort* 2) *Après s’être levé, le participant marche
 jusqu’à l’armoire où il va s’habiller* 3) *Le participant quitte la chambre et marche
 jusqu’à la cuisine pour préparer un repas* 4) *Le participant rejoint la table de la salle
 à manger, s’installe, mange en regardant la télévision* 5) *Le téléphone du participant
 sonne, il répond en quittant la table*



FIGURE 3.6 – Les capteurs utilisés et les propriétés d’objets auxquels ils se rattachent pour nourrir l’ontologie

le capteur de position indique quand l'humain est à côté de la cuisine. Cette dernière information permet au robot de savoir que l'humain est proche de lui. A l'aide d'un algorithme de patrouille proposé par la société Xaxxon, productrice du robot, il recherche dans la cuisine un humain et le suit. En fonction de son positionnement dans l'espace (odométrie), le robot sait où est l'humain et peut compléter les informations comme *humain est à côté de la Micro-onde* ou *humain est dans la cuisine*.

L'attention portée à la télévision est déterminée via la Kinect en utilisant l'orientation du visage obtenue via la bibliothèque openNi. En effet, lorsque l'axe des yeux est parallèle à celui de la télévision, on considère que le visage est orienté vers la télévision et donc que l'attention est portée sur cette dernière.

3.2 Évaluation des résultats et observations

Deux observateurs référents et 17 observateurs participants ont pris part à l'étude. Les réponses aux questionnaires sont reportées dans la figure 3.7.

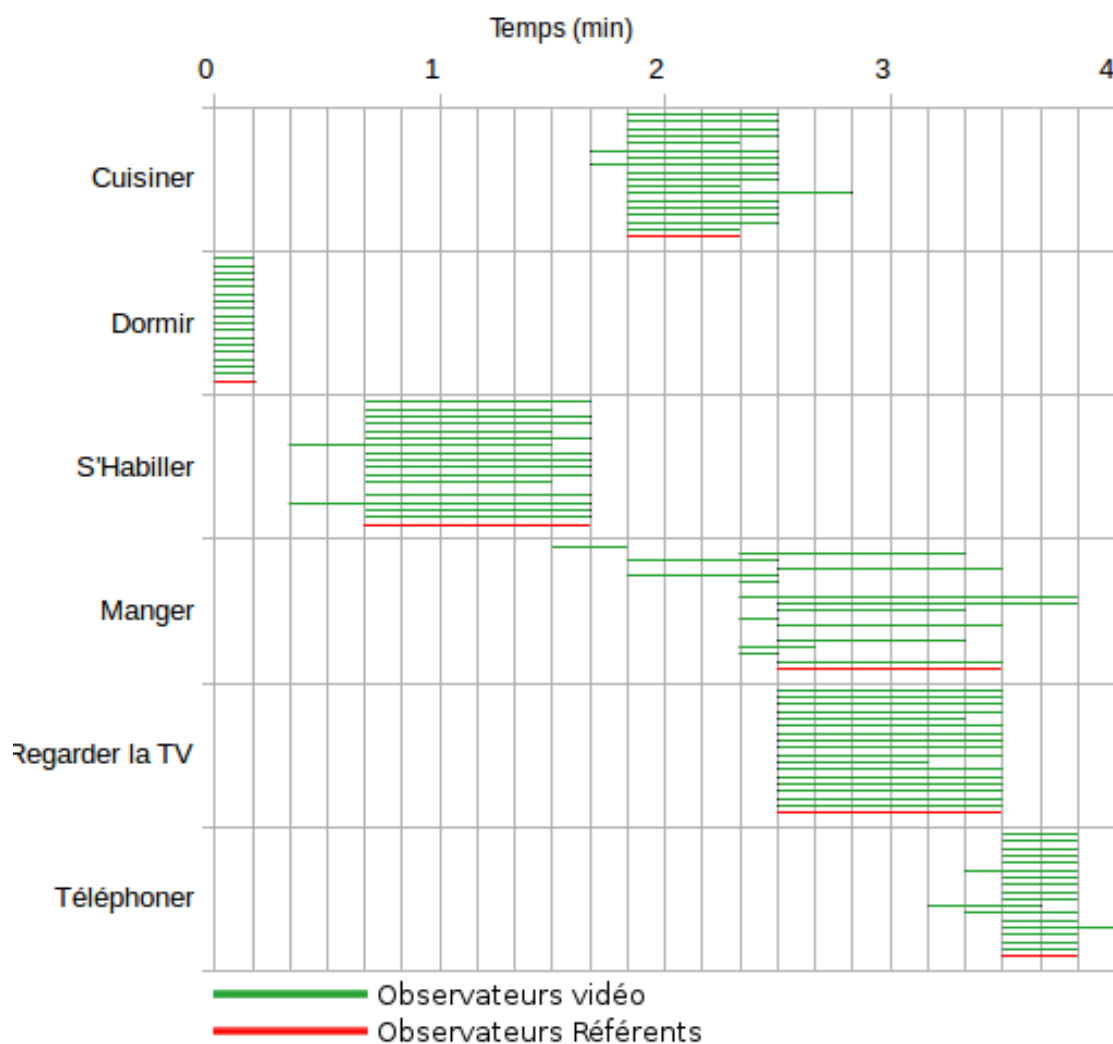


FIGURE 3.7 – Séquencement des activités d'après les observateurs référents et les observateurs participants

Exceptés pour l'activité *Dormir*, l'abstention d'un observateur sur un ou plusieurs horaires est représentée par l'absence d'indication sur l'intervalle temporelle de l'activité. Sur les informations des capteurs donnés aux OP, l'horaire de début de l'activité *Dormir* n'est pas explicitement positionné. L'interprétation du graphique 3.2 peut suggérer que l'activité a commencé avant le temps 00 :00. Les OP ont préféré ne pas se prononcer ne sachant si elle commençait effectivement à l'instant 0 ou en amont des enregistrements. Tous les résultats obtenus tiennent compte de cette abstention. Le tableau 3.2 résume l'interprétation de cette figure.

Interprétation des résultats en comparant les intervalles renseignés pour chaque activité

Afin de comparer les estimations des OP avec les OR, nous utilisons une méthode de comparaison par intervalles de temps. Cette dernière consiste à calculer la fréquence du nombre d'estimations correctes d'une activité pour chaque observateur durant l'intervalle de temps complet de l'activité :

Soient a et b respectivement l'horaire de début et l'horaire de fin d'une activité A , tous deux estimés par un observateur participant. Soient a_r et b_r respectivement les horaires de début et de fin de A indiqués par les observateurs référents.

On note $l(I)$ la fonction qui a un intervalle I associe sa longueur $l(I)$. De même $l(\bigcup_{k=1}^n I_k) = \sum_{k=1}^n l(I_k)$ pour I_1, I_2, \dots, I_n deux à deux disjoints.

Savoir quel est le pourcentage p_A de bonne estimation de l'activité A sur une période de temps revient à effectuer le calcul suivant :

$$p_A = \frac{l([a; b] \cap [a_r; b_r])}{l([a; b] \cup [a_r; b_r])} \times 100 \quad (3.1)$$

Dans le cas d'une abstention partielle (abstention sur l'horaire de début ou de fin d'une activité) d'un observateur participant, on considère $p_A = 0$.

La figure 3.8 illustre l'application de cette équation sur l'activité S'Habiller.

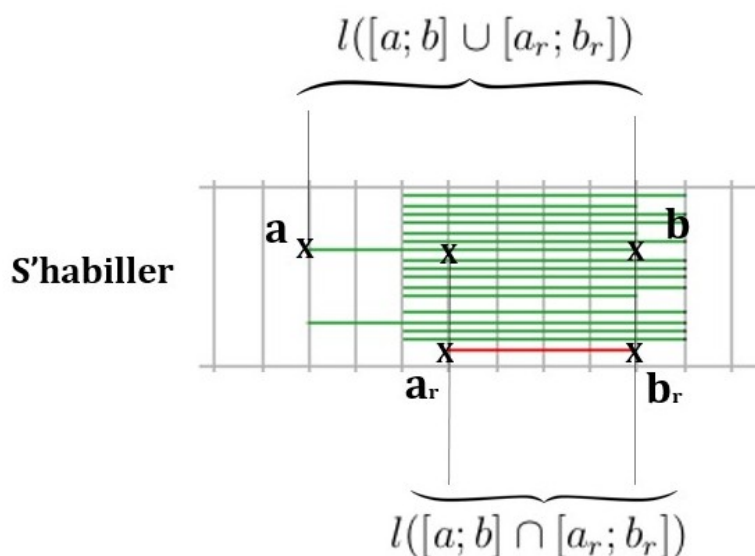


FIGURE 3.8 – Exemple d'application de l'équation 3.1 pour l'activité S'Habiller évaluée par un observateur référent et un observateur participant

Activités	$p_A(\%)$	Abstention (%)		Informations utilisées par au moins 70% des OP
		Début	Fin	
Cuisiner	76,17	0	0	humain estDans cuisine humain estProcheDe micro-onde
Dormir	82,35	17,65	0	Humain dortDans Lit humain estProcheDe lit
S’Habiller	63,56	5,88	5,88	humain estProcheDe lit humain estProcheDe Armoire
Manger	40,83	11,76	11,76	humain estDans salon humain estDans cuisine humain estProcheDe micro-onde humain regardeTVDans salon humain téléphoneDans salon
TV	97,06	0	0	humain regardeTVDans salon
Téléphoner	89,71	0	0	humain téléphoneDans salon
Moy. globale	74,95	5,88	2,94	

TABLEAU 3.2 – Tableau des abstentions et informations des évaluations des observateurs participants

Comparaison activité par activité

Grâce aux tableau 3.2 ainsi qu’à la figure 3.7, des conclusions peuvent être tirées sur les informations utilisées par les observateurs participants. Dans un premier temps, les activités directement détectée par les capteurs *Dormir*, *Regarder la télévision* et *téléphoner* semblent très bien reconnues par les OP. On observe que ces trois activités sont directement indexées sur les informations extraites des ontologies : *humain dortDans lit*, *humain regardeTVDans salon*, *humain téléphoneDans salon*. Concernant le début de l’activité *Dormir*, dans l’hypothèse où l’information précise du début du sommeil soit indiquée, nous pouvons estimer, d’après les résultats obtenus sur les activités *Téléphoner* et *Regarder la télévision*, que les observateurs participants s’appuieraient sur ces informations ; là où les observateurs référents se sont appuyés sur les mouvements du corps, résultant d’étirements et de la sortie du lit, pour indiquer la fin de l’activité du sommeil.

Aucune information des capteurs ne permet de déduire l’activité *Manger* et pourtant seulement 12% des OP se sont abstenus sur son estimation. Si on s’intéresse maintenant aux informations des capteurs utilisées par les OP pour déduire les horaires de début et de fin, on constate que les déductions faites s’appuient sur cinq informations là où pour chaque autre activité, les OP s’appuient sur une ou deux informations. L’emploi de ces informations ne permet donc pas de séquencer l’activité *Manger* en fonction des autres et donc d’estimer son horaire de début et de fin. D’autres informations sont donc nécessaires.

Les activités *S’Habiller* et *Cuisiner* sont comparables en termes de résultat et d’informations. Pour ces deux activités, un objet lié à l’activité est défini ; le micro-onde pour *Cuisiner* et l’armoire pour *s’Habiller*. Le moment où le participant s’approche d’un de ces objets et considéré comme étant le moment du début de l’activité correspondante. Lorsque le participant s’éloigne de cet objet, les observateurs participants considèrent que cette activité prend fin. Ainsi l’horaire de début de l’activité *Cuisiner* est correctement estimée avec environ 88% de bonnes estimations. Seulement l’horaire de fin de *Cuisiner* et les horaires de *s’Habiller* sont mal estimés.

Conclusion de la comparaison

Ces résultats démontrent que la reconnaissance des activités de la vie quotidienne est réalisable à travers notre système. En effet, la méthode utilisée permet de comparer le séquençement de chaque activité les unes par rapport aux autres. Une disparité est observée sur les activités *s’habiller* et *Manger* avec respectivement 64% et 41% de bonnes reconnaissances.

S’il est d’un intérêt limité de connaître les horaires précis de chaque activité, il reste néanmoins important d’avoir une estimation correcte. Détecter une anomalie sur l’exécution d’une activité comme par exemple le raccourcissement considérable du temps nécessaire pour manger ou au contraire l’augmentation importante du temps nécessaire pour se laver peut indiquer un problème chez la personne. Problème qu’il conviendrait de détecter et de remédier avec l’aide d’un professionnel de santé. En conclusion, l’utilisation des données anonymisées de la plate-forme permet à un humain de réaliser la reconnaissance et le séquençement des activités de la vie quotidienne.

3.3 Conclusion de l’étude et discussion

Suite à l’enregistrement d’activités de la vie courante réalisé par la plate-forme avec un participant humain, 17 observateurs participants ont analysé des informations délivrées par les ontologies. Leurs analyses ont consisté à retracer l’emploi du temps des activités réalisées par le participant en indiquant les horaires de début et de fin de chaque activité. Ces analyses sont comparées aux horaires indiqués par les observateurs référents ayant accès à l’intégralité des enregistrements vidéos et sonores. L’hypothèse de cette étude est que les informations sémantiques délivrées par la plate-forme aux OP permettent une reconnaissance fiable des activités de la vie courante.

En analysant le séquençement des activités fait par chaque OP on constate qu’il correspond au séquençement des activités des OR. En utilisant une méthode de correspondance entre les intervalles temporels de chaque activité estimés par les OP par rapport aux intervalles estimés par les OR, on obtient des résultats significatifs. Les informations de la plate-forme précisant directement le début et la fin d’une activité (*Dormir*, *Regarder la télévision* et *Téléphoner*) ont permis d’obtenir des observations similaires entre les observateurs. Cependant, on constate que le reste des informations de la plate-forme ne concernant pas directement des activités n’ont pas permis d’obtenir des résultats significatifs pour les activités restantes.

L’amélioration de ces résultats est amplement possible, notamment sur les activités *Manger*. Pour ce faire, de nouvelles informations pourraient être extraites par la plate-forme et délivrée aux OP, en utilisant par exemple le robot pour donner une information sur la proximité de l’humain par rapport à un objet lié à l’activité. *S’Habiller* est plus complexe. Comparer les habits d’un sujet humain reviendrait à le filmer en permanence à l’entrée ou sortie de pièces. Positionner des caméras dans un appartement est un sujet d’éthique délicat. Nous en employons mais la modularité de la plate-forme permet leur retrait à tout moment. Les informations liées directement à ces capteurs peuvent être récupérées d’une autre manière en employant, par exemple, les capteurs de mouvements.

Chapitre 4

Évaluation automatique de l'interaction Patient/Soignant au cours d'un soin

L'étude du comportement du soignant et du patient lors d'un soin médical est une approche importante pour analyser la qualité des soins. Des échelles d'évaluation renseignées par un observateur humain est l'approche classique pour analyser ces comportements et rendre état de l'interaction Patient/Soignant. Dans ce chapitre, nous proposons un système automatique permettant à un observateur humain d'évaluer une interaction entre un patient et un soignant. Ce système n'a pas pour vocation de remplacer l'évaluation humaine. En revanche, son utilisation permet de faciliter l'évaluation a posteriori du soin et en procurant un ensemble d'informations supplémentaires utiles pour l'évaluation.

4.1 Objectif, méthodologie et matériel de l'étude

Dans cette section nous détaillons la méthodologie nous ayant permis de mesurer les performances du système automatique ainsi que les outils informatiques et électroniques qui le composent. Pour cela nous avons réalisé deux études. La première nous permettant d'affirmer l'intérêt d'utiliser un système automatique, la seconde pour mesurer son efficacité.

4.1.1 Objectif, hypothèses et définitions

L'objectif de ce projet est de proposer un système automatique d'aide à l'évaluation de l'interaction entre un patient et un soignant lors d'un soin et d'évaluer ses performances. L'hypothèse principale que nous faisons est qu'un système automatique d'analyse du comportement permet à l'observateur l'employant, d'évaluer cette interaction et ce, sans altérer les résultats obtenus.

Deux autres hypothèses secondaires s'ajoutent à la première :

1. La présence de caméras lors d'un soin n'impacte pas le travail des soignants ni n'altère le comportement des patients et soignants,
2. L'utilisation d'enregistrements vidéos bruts a posteriori du soin permet d'évaluer le comportement.

Tout au long de ce chapitre nous souhaitons comparer les évaluations des différents types d'observateurs. Dans chaque situation, nous nous baserons sur l'évaluation *Observateur Référent* (OR). Les observateurs employant les enregistrements bruts de l'interaction

seront appelés *Observateurs Vidéos* (OV). Enfin l'observateur exploitant les informations délivrées par le système automatique est appelé *Observateur Système* (OS).

Des échelles de mesures permettant d'objectiver l'interaction entre le patient et le soignant lors d'un acte médical ont été développés par le Pr Belmin :

- L'échelle d'Evaluation du Comportement de la Personne Aidée au cours de l'intervention avec un Intervenant – ECPAI, voir l'échelle ECPAI dans l'annexe [A](#)
- L'Echelle de Comportement de l'Intervenant au cours d'une Interaction avec une Personne Aidée - ECIIPA, voir l'échelle ECIIPA dans l'annexe [B](#)

Ces échelles renseignent 10 comportements qu'ont eu le patient et le soignant durant leur interaction lors d'un soin. Une fois les échelles remplies, elles indiquent si le comportement de la personne observée a été globalement positif, neutre ou négatif pour l'interaction. Elles renseignent par exemple le taux de coopération du patient avec le soignant ou encore si ce dernier était calme et souriant.

4.1.2 Une étude, deux expériences

Lors de cette thèse, deux expériences distinctes ont été réalisées successivement afin de mettre en place et évaluer ce système automatique. Pour chacune des deux étapes, à minima un comité d'éthique a validé les protocoles en amont.

Mise en place et première évaluation du système automatique dans un environnement contrôlé

La première expérience vise à mesurer les conséquences d'une évaluation du comportement en utilisant des enregistrements au lieu d'être présent. En effet, les échelles ECPAI et ECIIPA ont toujours été remplies par un observateur présent. Afin de pouvoir réaliser l'évaluation à distance, il est nécessaire de comprendre les conditions optimales permettant à l'observateur de réaliser son analyse. Nous avons donc filmé l'interaction entre 16 faux patients et 5 faux soignants en utilisant 3 caméras. Ces enregistrements ont ensuite été montrés à un observateur habitué à remplir les échelles d'évaluation. Les retours de cet observateur ont permis de mettre en place le protocole pour l'expérience suivante. De plus, ses évaluations des interactions ont permis de calibrer les algorithmes d'analyse automatique du comportement.

Étude en situation réelle et premières mesures d'efficacité

Une fois les algorithmes étalonnés, la seconde partie de l'expérience a été réalisée en situation réelle à l'hôpital. L'objectif ici est de mesurer les performances du système à informer un observateur afin que ce dernier évalue l'interaction. L'interaction entre 3 soignants et 8 patients a été enregistrée comme pilote pour étude. 4 observateurs ont été recrutés afin de remplir les échelles ECPAI et ECIIPA : un observateur référent, deux vidéos et un système. Les différentes échelles ainsi obtenues ont été comparées entre elles afin de déterminer si : 1) l'utilisation des enregistrements pouvait mener à une évaluation satisfaisante des comportements ; 2) l'utilisation du système pour aider l'évaluation a elle aussi permis de réaliser une évaluation satisfaisante par l'OS.

4.1.3 Spécifications techniques

Le système que nous avons mis en place pour ces expériences est un dérivé du système expliqué dans le chapitre [2](#).

Matériel employé

Tout au long de ces études, seul le système ambiant a été utilisé. Nous cherchons ici à éviter qu'un observateur soit présent lors du soin et nous tenons à vérifier l'hypothèse que les caméras ne gênent pas le travail des soignants. Employer un robot mobile ne pouvait garantir ce point en cas d'urgence. Ainsi, la plate-forme présentée dans le chapitre 2 étant modulable, le robot peut donc être retiré sans déranger le fonctionnement du système ambiant qui reste indépendant.

Durant les enregistrements réalisés dans la première étude, seules 2 kinects ont été utilisées pour le système. Une caméra grand angle enregistrait l'intégralité de la scène. Cette dernière n'était pas reliée au système automatique, ses enregistrements ont été utilisés par l'observateur référent pour évaluer les comportements. Le schéma de la figure 4.1 montre le positionnement des caméras dans la salle des enregistrements en fonction des différentes étapes du protocole. La figure 4.2 montre les différents points de vue des deux kinects ainsi que de la caméra grand angle.

Pour la seconde étude, le système ambiant est composé uniquement d'une kinect positionnée sur un chariot médical mobile (voir la figure 4.4). Le chariot médical n'est pas autonome et nécessite l'intervention humaine pour être déplacé.

Bibliothèques informatiques et algorithmes employés

En utilisant la Kinect comme capteur principal, trois types de données étaient récoltées durant les enregistrements : Le son de l'environnement, la vidéo en format RVB, la profondeur de champs de la caméra.

En utilisant la "OpenNI Kinect Library", le squelette d'un humain peut être extrait donnant accès à coordonnées en 3 dimensions des mains et de la tête ainsi que l'orientation de la tête. La détection du squelette d'un humain en utilisant cette librairie nécessite la prise d'une position appelée "psi pose" ou position psi. Cependant, une option permettant la calibration automatique des kinects a été utilisée. Après plusieurs essais, il a été mis en évidence que le système nécessitait un certains temps, pouvant aller jusqu'à plusieurs minutes, afin de pouvoir afficher le squelette. Pour être certain d'obtenir le squelette, le scénario incluait la prise de la position *psi*, "justifiée" dans le scénario sous la forme d'un exercice que le soignant fait faire au patient.

A partir de la caméra optique, la détection du corps humain, sa tête et le sourire sont extraits en utilisant la méthode proposée par Viola et Jones dans l'article [Viola and Jones, 2001].

La localisation de son a été rendue possible en utilisant le logiciel HARK [Nakadai et al., 2008]. On considère qu'un participant a parlé en fonction de l'amplitude sonore venant des sons situés au niveau de la position du soignant, si ceux-ci sont plus élevés que les autres, nous considérons qu'il parle. En effet, les bruits engendrés par l'interaction lors du bandage sont négligeables en terme d'amplitude et de durée par rapport au volume sonore induit par la communication.

La quantité de mouvement (QoM) réalisée par un utilisateur peut être employée afin de mesurer un niveau de stress [Aigrain et al., 2016]. La kinect donnant des données sur les jambes instables lorsque le participant est assis, seule la quantité de mouvement de la partie supérieure du squelette, couplée à la quantité de pixels ayant été modifiés entre deux images (iQoM) a été utilisée pour cet algorithme.

En utilisant les méthodes et algorithmes présentés ci-dessus, le système automatique était chargé de travailler sur les comportements, décrits par les échelles, à observer suivants :

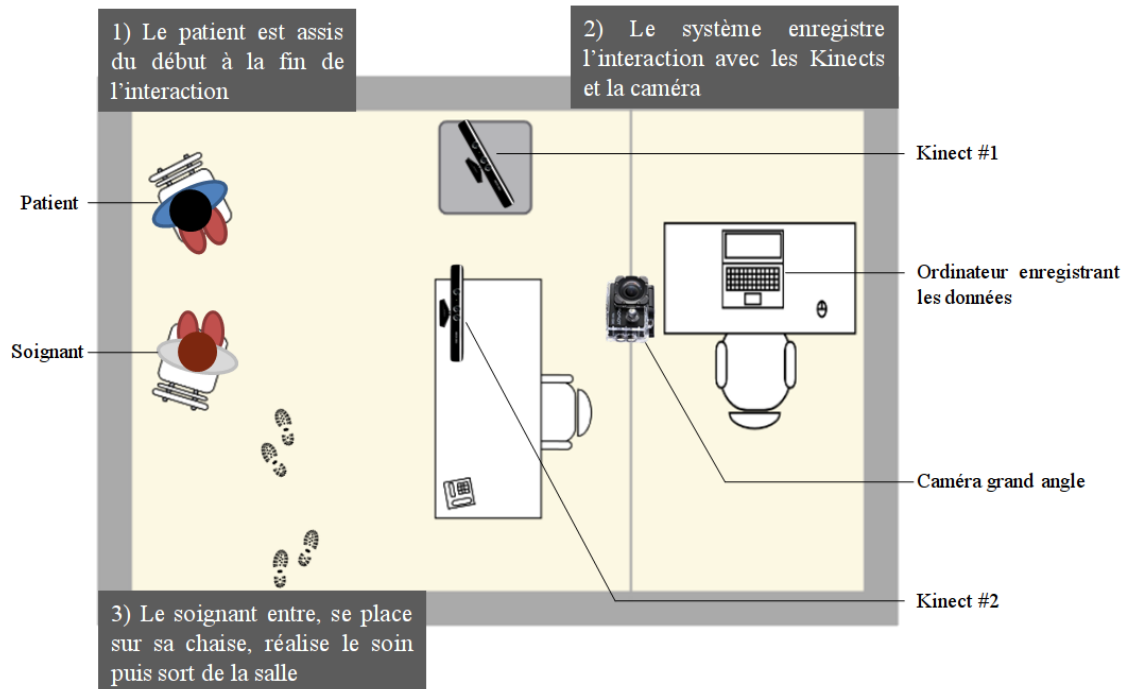


FIGURE 4.1 – Schéma explicatif de la première expérience

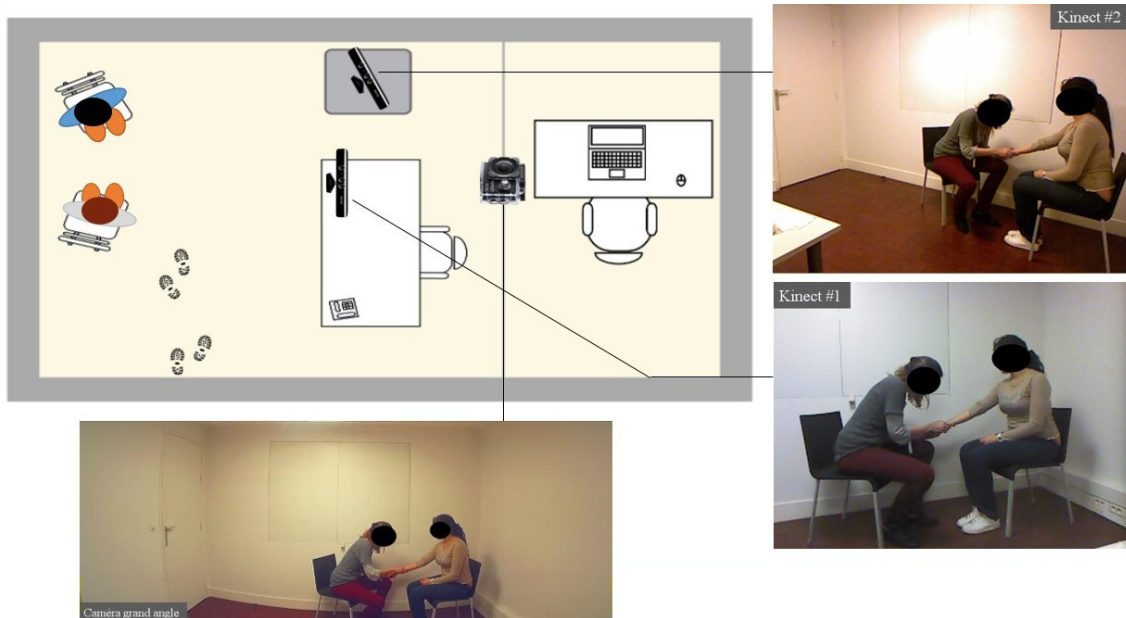


FIGURE 4.2 – Point de vue des différentes kinects et de la caméra grand angle

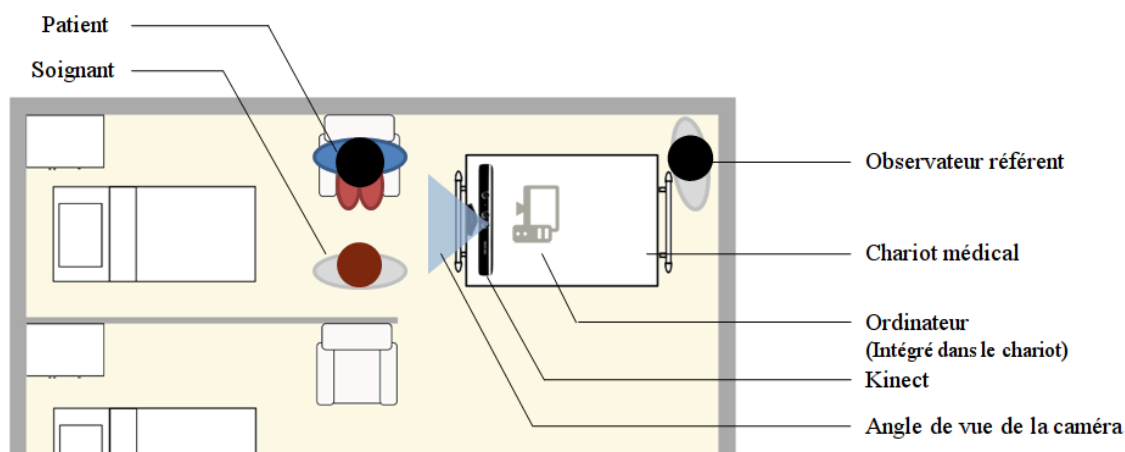


FIGURE 4.3 – Séquençage des activités d’après les observateurs référents et les observateurs participants

Pour l’échelle ECIIPA (voir annexe B)

1. *Au départ, a cherché à capter le regard de la personne aidée* est déterminé par l’orientation de la tête lorsque le soignant entre dans la pièce et s’approche du patient (Openni Kinect Library),
2. *Au départ, a cherché à établir un contact verbal avec la personne aidée*, est déterminé en cherchant à analyser le moment où l’intervenant a parlé en premier (Hark sound localization),
3. *A touché la personne aidée après avoir établi préalablement une relation*, est déterminé en cherchant à analyser si l’intervenant a parlé avant de toucher la personne (Openni Kinect Library & Hark sound localization),
4. *Semble calme et détendu* est obtenu en mesurant la quantité de mouvement [Aigrain et al., 2016],
5. *Est souriant et exprime de la bienveillance envers la personne aidée* est déterminé en détectant le sourire [Viola and Jones, 2001],
6. *A veillé à maintenir le contact durant toute l’interaction* est déterminé en fonction du temps que l’intervenant a passé à parler avec la personne aidée par rapport au temps de l’interaction (Hark sound localization).

Aucune étude sur la linguistique employée n’est réalisée. En conséquence les autres comportements ne sont pas analysés.

Pour l’échelle ECPAI (voir annexe A)

1. *Est attentive à l’intervenant ou le suit des yeux* est déterminé en fonction de l’orientation de la tête lors de l’interaction (Openni Kinect Library)
2. *Cherche à communiquer verbalement avec l’intervenant* est déterminé en fonction du temps de parole de la personne aidée par rapport au temps de l’interaction (Hark sound localization)
3. *Semble calme* est mesuré à l’aide de la mesure de la quantité de mouvement [Aigrain et al., 2016]
4. *Aide à l’action de l’intervenant* est mesuré par le fait que la personne aidée soit touchée ou non (Openni Kinect Library),



FIGURE 4.4 – Photographie illustrant la mise en place du chariot médical. La kinect positionnée sur le plateau principal du chariot permet de filmer le soignant lorsqu’il rentre et sort de la chambre. En réorientant le chariot durant le soin, la kinect filme l’interaction. Les tiroirs et le plateau peuvent être utiliser pour poser du matériel de soin

Il est a noter que malgré la possibilité de faire varier l’annotation d’un comportement sur une échelle entre 5 degrés d’estimations, certains comportements sont détectés de manière binaire "a eu lieu" ou "n’a pas eu lieu". Ainsi les comportements suivants seront considérés comme "Comportement A/B de façon nette" en fonction des observations faites par le système :

- Au départ, a cherché à capter le regard de la personne aidée,
- Au départ, a cherché à établir un contact verbal avec la personne aidée,
- A touché la personne aidée après avoir établi préalablement une relation,
- A veillé à maintenir le contact durant toute l’interaction,
- Aide à l’action de l’intervenant.

Les informations ainsi extraites par le système nourrissent les ontologies.

4.2 Expérience réalisée au sein de l'INSEAD Behavioural Lab

L'étude réalisée au sein de l'INSEAD-Sorbonne Université Behavioural Lab ¹ consiste à filmer des interactions entre deux participants. Pour chaque participant un rôle de patient ou de soignant lui est attribué. Durant un scénario qui leur est expliqué en amont, leurs comportements et interactions seront enregistrés par la plate-forme décrite dans la sous section 4.1.3. A l'issue des enregistrements, un questionnaire comprenant les échelles ECPAI et ECIIPA ainsi qu'une question sur leur ressenti par rapport aux caméras leur est donné afin de les remplir.

Dans un deuxième temps, les interactions sont évaluées par un observateur habitué à utiliser les échelles. Cet observateur est considéré comme l'observateur référent de l'étude. Toutes les échelles sont alors comparées deux à deux afin de déterminer si l'évaluation faite à partir d'enregistrement vidéo est satisfaisante ou donne des résultats disparates.

Enfin, la plate-forme automatique est utilisée pour extraire des informations. Ces informations sont directement et automatiquement utilisées pour remplir une partie des échelles. Une fois cette étape passée, ces dernières sont comparées avec celles de l'observateur référent.

4.2.1 Protocole

Scénario

Le scénario se déroule de la manière suivante :

1. Il débute dès lors que le soignant rentre dans la pièce et que l'interaction commence. Comme indiqué dans la figure 4.1, le patient est assis sur une chaise pendant que les caméras enregistrent.
2. Le soignant prend immédiatement position sur la chaise en face du patient et lui pose quelques questions sur le sujet de sa venue. Quelques observations sont faites puis ce dernier s'applique à poser un bandage sur l'avant bras de la personne.
3. Enfin, quelques recommandations orales sont dispensées avant que le soignant quitte la salle.

La durée moyenne des enregistrements est d'approximativement 3 minutes. Ce scénario a été commun à tous. Des consignes sont données aux participants pour le respecter mais aucune ne demande de réaliser des actions supplémentaires. En revanche, en fonction du comportement du patient, seule la pose du bandage pouvait être optionnelle. Le sujet de discussion devait rester approprié au contexte mais aucune importance n'a été donné sur la pertinence du contenu qu'il soit d'ordre médical ou simplement social.

Les participants, du recrutement aux consignes

Le recrutement des participants et l'étude complète se sont réalisés au sein des locaux de l'INSEAD Behavioural Lab. Aucune consigne d'âge, de genre ou de particularité quelconque n'a été donné concernant le recrutement des participants jouant le rôle des patients. Le but de ces travaux n'est pas d'étudier le comportement en fonction de paramètres physiques ou sociaux mais au sein de l'interaction. En conséquence, et conformément au Règlement Général sur la Protection des Données (RGDP) aucune information personnelle n'a été enregistrée. 16 participants jouant le rôle du patient ont été recrutés.

Concernant le rôle de soignant, le scénario prévoyant des discussions sur le sujet médical

¹ <https://www.insead.edu/centres/insead-sorbonne-universite-lab-fr>

et la pose d'un bandage au niveau de l'avant bras droit du patient, il a été décidé de ne recruter que des personnes appartenant, ou ayant, appartenu au milieu médical. Ceci afin de faciliter la mise en contexte et l'immersion des participants durant le scénario. 5 participants ont été recrutés pour interpréter le rôle du soignant, parmi eux figuraient des proches et personnes ayant déjà réalisés des études dans le cadre d'autres recherches réalisées au sein de l'INSEAD Behavioural Lab.

L'accueil des participants était individuel, de manière à ce qu'aucun d'entre eux ne se connaissent et n'ait déjà interagi au préalable avec l'autre avant l'interprétation du scénario. Après une courte introduction, le scénario était détaillé et le rôle de chacun était clarifié. Afin d'obtenir de la variété sur les comportements à observer, quelques consignes étaient données aux participants en suivant des points spécifiques sur les échelles. Il était demandé au patient de rester sur sa chaise durant toute la durée de l'interaction. En revanche, si celui-ci le souhaitait, il pouvait refuser le soin en se tournant sur sa chaise pour ne plus regarder le soignant, cacher ou protéger son avant bras. Cependant, aucune information précise concernant les échelles n'était donnée : "*Pourriez-vous chercher à énormément communiquer avec le soignant ?*" ou encore "*Ne cherchez pas le regard du patient lorsque vous entrez dans la salle, faites en sorte de réaliser le scénario le plus rapidement possible*". Aucune consigne supplémentaire n'a été donnée concernant une quelconque attention supplémentaire à porter au partenaire de l'interaction.

A l'issu des enregistrements, un questionnaire comprenant la question "*Sur une échelle de 1 à 5, à combien estimeriez vous l'attention que vous avez porté aux caméras ? 1 étant Je n'y ai pas prêté attention, 5 étant j'y ai porté beaucoup d'attention.*" ainsi que les échelles ECPAI et ECIPA était donné aux participants. Ces derniers étaient séparés dans deux salles différentes et y répondaient de manière individuelle. Enfin, un compte rendu était réalisé avec le patient afin d'expliquer les objectifs de cette étude et l'utilisation que nous allions faire des enregistrements. Le soignant enregistrant plusieurs sessions successives, ce même compte rendu n'était réalisé qu'à la fin de ses sessions. Chaque participant repartait avec une lettre résumant les objectifs de cette étude (voir l'annexe D).

4.2.2 Les évaluations

Sur 21 participants, seuls 2 ont renseigné avoir porté "légèrement"² attention aux caméras lors des enregistrements. Les 19 autres ont répondu ne pas avoir porté du tout attention aux caméras. Cependant 3 d'entre eux ont complété à l'oral qu'ils étaient plus occupés à appliquer les éventuelles recommandations faites sur leur comportement. Ces premiers résultats tendent à indiquer que la présence de caméras lors d'une interaction scénarisée n'attire pas l'attention des participants.

Sur 16 enregistrements, 15 ont pu être utilisés durant l'étude. Des données corrompues ont rendu inutilisables le 16^e enregistrement.

Dans les prochaines sous sections, nous comparons deux à deux les échelles obtenues. Par soucis de clarté, seuls les tableaux de résultats sont présents dans cette section. Vous retrouverez les résultats des comparatifs des échelles dans les annexes C.

² Ce qui correspond à 2 sur 5 sur l'échelle de la question posée dans le questionnaire

4.2.3 Résultats

Pour chaque comparaison entre deux échelles, nous avons cherchés à connaître le nombre de case d'écart entre les deux évaluations, voir figure 4.5. D'un individu à l'autre, l'observation et l'analyse de certains détails dans le comportement peut varier. Nous n'accorderons pas d'importance à une variation d'une case d'écart. Ainsi, si un observateur évalue que le soignant a souri de façon "très nette" durant toute l'interaction, tant qu'un second observateur évalue qu'il a souri de façon "nette" uniquement, nous ne considérons pas cela comme deux évaluations contradictoires (voir les échelles A et B pour plus de détails sur les différents comportements). En revanche, un écart plus grand entre deux notations ou l'absence d'une évaluation sont considérés comme deux évaluations différentes pour le même comportement. En effet, cette différence d'écart peut correspondre à avoir un comportement *confiant* de façon "nette" à un comportement *inquiet* de façon "nette". Nous faisons la distinction entre les deux échelles. Ainsi ces écarts sont déterminés en fonction de chaque comportement du patient puis, du soignant. Les tableaux 4.1, 4.2, 4.3 et 4.4 sont exprimés en fonction des pourcentages pour chaque écart.

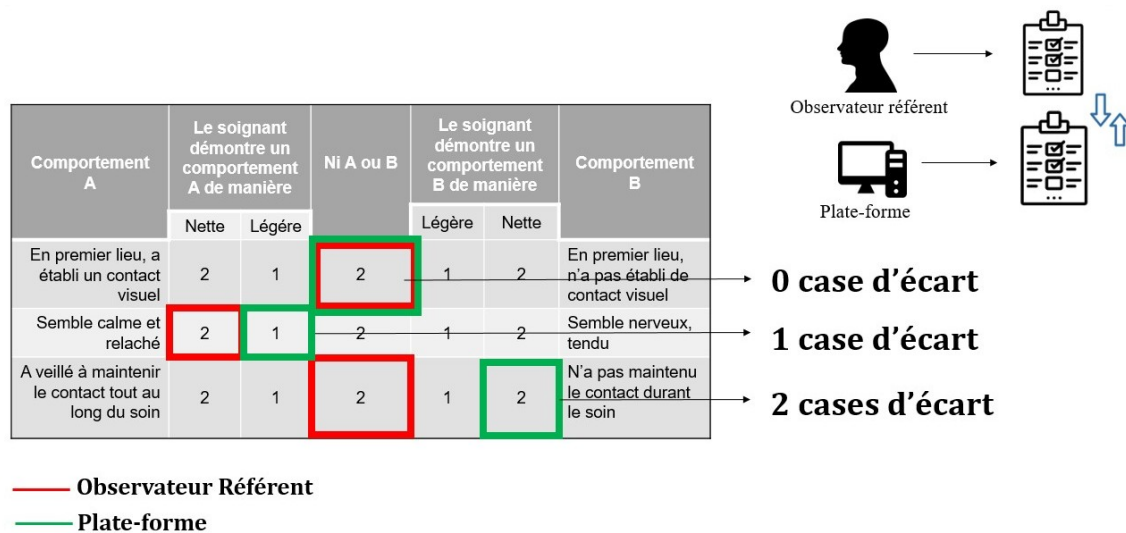


FIGURE 4.5 – Exemple de comparaison d'évaluation du comportement d'un soignant réalisée par un observateur référent et la plate-forme

Comparaison des évaluations entre le patient et le soignant

En premier lieu, les grilles d'évaluations annotées par le participant jouant le rôle du patient et celles annotées par le participant jouant le rôle du soignant ont été comparées. Les grilles sont accessibles dans l'annexe C.1). Le tableau 4.1 résume cette comparaison. Il est à préciser que les participants ne connaissaient pas les échelles et n'ont eu aucune formation spécifique avant de les remplir.

Pour chaque dyade, 46% des échelles ECIIPA correspondent exactement et 40% pour les échelles ECPAI. Ces résultats passent respectivement à 78% et 70% en prenant en compte une seule case d'écart dans l'évaluation. Concernant l'échelle ECIIPA, les principales divergences viennent de la question concernant la demande de consentement avant la réalisation d'un geste. Les principales divergences concernant les échelles ECPAI correspondent à l'expression de sympathie ou d'hostilité.

Ces résultats démontrent que les évaluations sont relativement similaires. On pourra attribuer les désaccords entre les participants au fait qu'ils étaient à la fois observateurs

Nombre de cases d'écart entre les évaluations	0	1	2	3	4
Echelle ECPAI (%)	46	32	13	6	3
Echelle ECIIPA (%)	40	30	22	7	1

TABLEAU 4.1 – Résultats de la comparaison des échelles annotées par les observateurs vidéos et l'observateur référent

et acteurs brouillant ainsi leur propre perception de soi [Stone et al., 1999, Aigrain et al., 2016]. De plus, il ne leur a pas été demandé en amont porter une attention supplémentaire aux comportements de leur partenaire d'interaction.

On remarquera que 56% des discordances (écart supérieur ou égal à 2 cases) entre les réponses des deux participants sont réalisées par 5 dyades, les numéros 6, 7, 8, 12 et 14. Ces 5 interactions comprenant 4 soignants différents, elles ne leur sont pas imputables.

Comparaison des évaluations entre le patient, le soignant et l'observateur référent

Les tableaux 4.2 et 4.3 résument la comparaison faite sur les grilles d'évaluations annotées par le participant jouant le rôle du patient et celles du soignant deux à deux avec l'observateur référent (voir annexes C.3 et C.2).

Nombre de cases d'écart entre les évaluations	0	1	2	3	4
Echelle ECIIPA (%)	42	38	14	5	1
Echelle ECPAI (%)	39	34	21	5	1

TABLEAU 4.2 – Résultats de la comparaison des échelles annotées par l'OR et le patient

Nombre de cases d'écart entre les évaluations	0	1	2	3	4
Echelle ECIIPA (%)	52	34	13	1	0
Echelle ECPAI (%)	45	31	20	3	1

TABLEAU 4.3 – Résultats de la comparaison des échelles annotées par l'OR et le soignant

Les résultats obtenus lors de la comparaison patient/OR sont semblables à ceux obtenus par la comparaison soignant/OR, respectivement 80% et 86% pour les ECIIPA et 73% et 76% pour les ECPAI. Il n'y a donc pas d'affinité d'évaluation particulièrement remarquable entre l'un ou l'autre des participants avec l'OR. En revanche, ces résultats indiquent qu'un observateur et acteur de la scène, qui plus est non formé à l'annotation de ces échelles, évalue de manière similaire par rapport à un observateur extérieur habitué à évaluer. La faible quantité d'échantillons ne permet pas de conclure définitivement sur cette hypothèse. 48% des désaccords entre l'OV et le patient se font sur dyades numérotées 1, 6, 7 et 11. 50% des désaccords entre le soignant et l'OV se font sur les dyades numérotées 8, 11, 12 et 13.

En revanche, suite à ces résultats, ainsi qu'aux retours de l'OR, évaluer l'interaction du patient et du soignant à travers l'utilisation de vidéos est réalisable ce qui confirme notre hypothèse émise dans la sous section 4.1.1.

Comparaison des évaluations entre l’observateur référent et le système automatique

L’annexe C.4 montre les évaluations réalisées par le système. On remarque, tout d’abord, hormis les sourires parfois peu détectés, les comportements du soignant observés et analysés automatiquement sont identiques : La recherche du regard, l’établissement de la relation, la demande d’approbation, le maintien de la discussion, etc. Le tableau 4.4 résume la comparaison faite sur les grilles d’évaluations annotées par l’observateur référent et celles annotées par le système automatique (voir annexe C.5).

Nombre de cases d’écart entre les évaluations	0	1	2	3	4
Echelle ECIIPA (%)	77	21	2	0	0
Echelle ECPAI (%)	55	30	10	5	0

TABLEAU 4.4 – Résultats de la comparaison des échelles annotées par l’OR et le système automatique

Après avoir enregistré l’interaction, le système automatique a été chargé d’analyser les interactions en utilisant les algorithmes présentés dans la sous section 4.1.3. Dans le tableau 4.4, seules les 5 cases de l’échelle ECIIPA et les 4 cases de l’échelle ECPAI que peut analyser la plate-forme sont comparées. Sans surprise, les résultats démontrent que l’étalonnage des algorithmes correspondent aux évaluations de l’OR. Plusieurs essais ont été réalisés par dichotomie jusqu’à obtenir ces évaluations jugées suffisantes pour considérer que le système évalue d’une manière proche à celle de l’OR.

4.2.4 Conclusion

Nous avons mis en place un protocole permettant d’étalonner notre plate-forme afin qu’elle soit en mesure de réaliser une évaluation de l’interaction entre un patient et un soignant. Ce protocole s’est déroulé selon 3 étapes. La première consistait à réaliser des enregistrements d’interactions et la faire évaluer par les acteurs eux-mêmes ainsi que par un observateur extérieur habitué à employer ces échelles et donc considéré comme référent. Cet observateur a employé des enregistrements vidéos afin de réaliser ses évaluations. La comparaison de ses échelles avec celles des participants démontre que leurs trois analyses sont relativement similaires (de l’ordre de 80% de corrélations pour les échelles ECIIPA et 75% pour les échelles ECPAI).

Les dyades numérotées 6, 7 et 8 augmentent à elles seules le taux de désaccord lors des comparaisons entre le patient et le soignant et entre le l’OV et le patient (dyades 6 et 7) et l’OV et le soignant (dyade 8). Les interactions 6 et 7 ont été réalisées par le même soignant, seulement la comparaison entre les évaluations du soignant et de l’observateur et l’OV concordent à hauteur respectivement de 20% et 15% soit respectivement 4 et 3 désaccords sur les 2 échelles. Les principaux désaccords observés lors de ces comparaisons ne sont donc pas imputables à un seul participant. Augmenter le nombre d’observateurs permettrait uniquement de conclure sur la capacité que ces participants ont pour annoter les échelles. Nous ne les avons donc pas retiré ces données de notre étude.

A la question de la présence de caméras lors de l’interaction, les participants ont déclarés ne pas avoir porté leur attention dessus. Cependant, certains ont spécifié qu’ils étaient plus occupés à appliquer les recommandations faites sur leur comportement. A la vue de ces réponses, nous ne souhaitons pas conclure de manière définitive sur une absence d’altération du comportement lié à la présence de caméras.

Enfin, nous avons comparé l’analyse de l’interaction réalisée par la plate-forme a celle de l’observateur référent. Les résultats indiquent que sur les annotations faites par le système,

une grande corrélation existe avec celles de l'OR. Le système est donc en mesure d'évaluer de manière relativement similaire certains comportements définis du patient et du soignant.

Dans les sections précédentes, nous avons analysés les évaluations annotées par chaque participant, un référent vidéo et le système. Les grilles d'évaluations du soignant n'offrent pas une grande diversité parmi un certain nombre de réponses, notamment celles concernant des faits mesurables de façon simple : La recherche du regard, l'établissement de la relation, la demande d'approbation avant de toucher la personne. Le nombre d'échantillons de soignants utilisés pour réaliser cette analyse ne nous permet pas de conclure sur la mise en évidence d'un comportement commun à tous.

En revanche, le comportement des patients offrent une diversité plus large. Ceci permet d'observer une plus grande variation dans les évaluations proposées par le système automatique. Démontrant sa capacité à analyser une partie de leurs comportements.

En revanche nous pouvons nous interroger sur la pertinence de la mesure de certains comportements du soignant dans notre étude :

Tout d'abord, le patient était situé face à la porte d'entrée. Ainsi lorsque le soignant entrait dans la pièce, il pouvait facilement entrer en contact visuel avec ce dernier. Ceci justifie l'absence de variations concrètes sur ce comportement de l'échelle ECIIPA.

De plus, les premiers échanges entre les participants ont été des présentations et des questions sur une douleur hypothétique justifiant le bandage et ce, sans que l'influence des consignes ne semblent affecter ces comportements. Suite aux comptes rendus réalisés avec les soignants, ces derniers ont admis avoir eu du mal à jouer un comportement réellement négatif lorsque demandé. Et ce, malgré la prodigation de conseils en lien avec les échelles comme "*Ne regardez pas le patient immédiatement, ne demandez pas la permission pour prendre le bras, etc.*". Le système n'est pas en mesure actuellement de faire la différence entre une discussion sociale ou une demande explicite d'autorisation pour toucher la personne aidée. Or, si certains conseils proposés tiennent plus de la linguistique, le système ne pourra pas déceler de variation à annoter sur les échelles.

Malgré tout, le système automatique obtient des résultats similaires par rapport aux évaluations de l'observateur référent. Cependant le manque de variation dans les comportements du soignant ne permettent pas de dire si le système est en mesure d'analyser efficacement les comportements inverses des soignants.

Une étude secondaire pourrait s'intéresser à l'intérêt de l'évaluation de ces comportements qui semblent être automatiquement positifs chez les personnes appartenant au corps médical et, à titre plus général, présents chez toutes les personnes oeuvrant dans des oeuvres sociales.

4.3 Expérience réalisée au sein de l'hôpital Charles Foix

Suite à l'étude réalisée dans un environnement entièrement contrôlé au sein de l'INSEAD-Sorbonne Université Behavioural Lab, nous avons réalisé sa continuité à l'hôpital Charles Foix. L'interaction de soin était cette fois-ci, enregistrée et analysée en conditions réelles. L'objectif de cette étude est de mesurer la performance du système automatique à évaluer l'interaction entre un patient et un soignant lors d'une prise de sang. L'étude a porté sur 8 personnes hospitalisées en gériatrie devant avoir une prise de sang dans le cadre de leurs soins médicaux, et acceptant de participer ainsi que 3 infirmiers chargés de réaliser cette prise de sang. Le schéma 4.6 résume les différentes étapes du protocole :

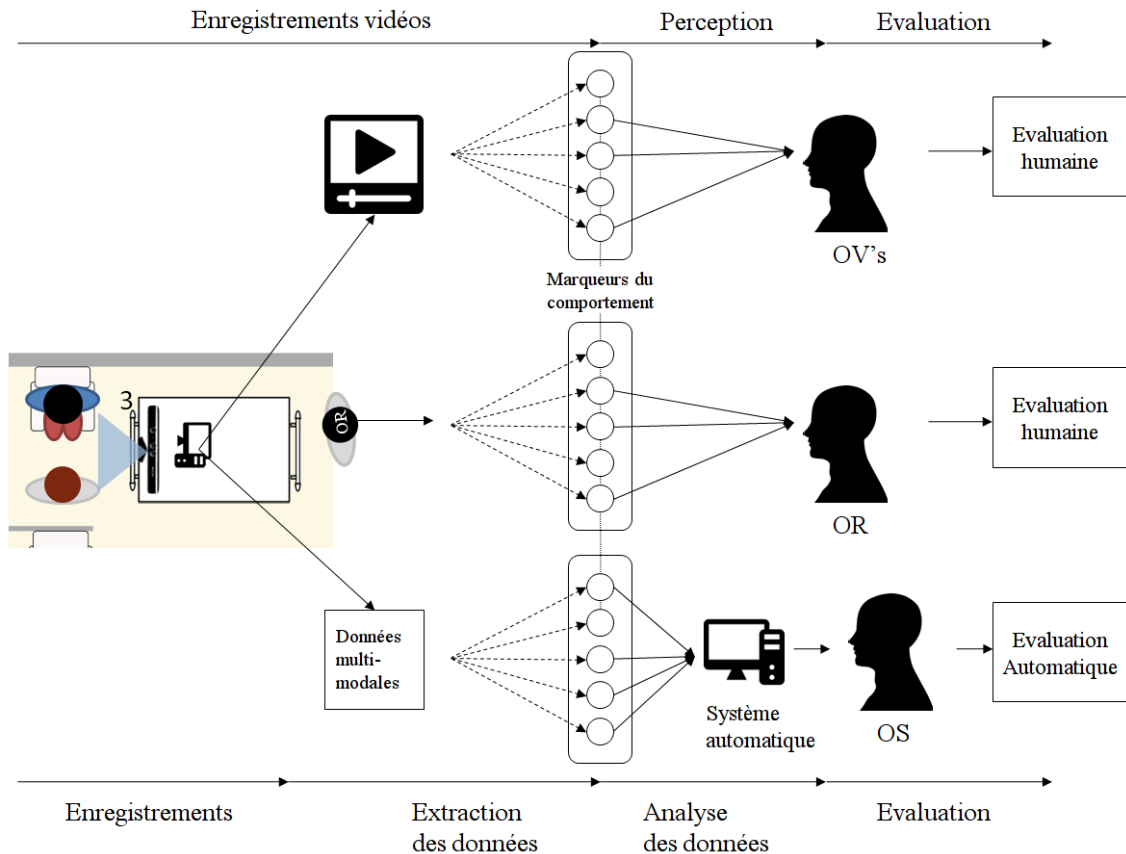


FIGURE 4.6 – Schéma des évaluations des comportements du patient et du soignant au cours de cette étude. L’interaction patient-soignant est évaluée de trois manières différentes. La première est réalisée par un observateur présent dans la chambre (OR) durant le soin. La deuxième analyse est réalisée par des observateurs utilisant les enregistrements bruts de la plate-forme (OV). Enfin, la dernière analyse est réalisée à posteriori par un observateur (OS) s’appuyant sur les informations extraites par le système automatique

La première consiste à enregistrer, à l’aide d’une caméra Kinect, l’interaction entre le patient et le soignant durant cette prise de sang. De plus, durant cet enregistrement, un observateur sera présent dans la chambre et analysera le comportement. Cet observateur habitué à annoter les échelles ECPAI et ECIIPA (voir respectivement les annexes A et B) est considéré comme étant l’Observateur Référent (OR).

Ensuite, des Observateurs Vidéos (OV) utiliseront les enregistrements audio et vidéo afin de remplir les grilles d’évaluation. Leurs grilles seront comparées à celle de l’OR afin de déterminer si l’évaluation est réalisable à partir du point de vue du système.

Enfin, dans une dernière étape, les enregistrements réalisés durant le soin seront analysés par un système automatique qui annotera les vidéos d’informations complémentaires. Un observateur extérieur évaluera alors l’interaction Patient/Soignant à partir de ces informations.

Les participants

L'étude a porté sur 8 patients³ hospitalisés en gériatrie, devant avoir une prise de sang programmée dans le cadre de leurs soins médicaux les jours de réalisation de l'étude. Le nombre de sujets nécessaires a été établi de façon empirique dans la mesure où il s'agit d'une étude pilote visant à établir la faisabilité et la preuve du concept. Afin de réaliser ces prises de sangs, 3 infirmier(e)s ont été inclus dans l'étude.

Chaque participant a été recruté après avoir vérifié que les critères d'inclusions et d'exclusions suivants étaient respectés :

Pour les patients, sont des critères d'inclusion :

- Homme ou femme hospitalisé(e) dans le service de gériatrie à orientation cardiologique et neurologique de l'hôpital gériatrique Charles Foix,
- nécessitant une prise de sang dans le cadre du soin hospitalier,
- pouvant être mis en fauteuil durant la prise de sang,
- capable de lire et comprendre la lettre d'informations sur les conditions de l'étude,
- capable de comprendre et donner un accord oral de participation à l'étude,
- capable de lire, comprendre et signer l'autorisation audiovisuelle

Sont des critères de non inclusion :

- Patient faisant l'objet d'une mesure de protection juridique,
- participant à un autre protocole d'étude.

Pour les soignants sont des critères d'inclusion :

- Infirmier(ère) travaillant dans le service de gériatrie à orientation cardiologique et neurologique de l'hôpital gériatrique Charles Foix,
- capable de lire et comprendre la lettre d'informations sur les conditions de l'étude,
- capable de comprendre et donner un accord oral de participation à l'étude,
- capable de lire, comprendre et signer l'autorisation audiovisuelle.

Sont des critères de non inclusion :

- Soignant participant à un autre protocole d'étude.

Les participants de l'étude ont été recrutés dans le cadre des visites médicales journalières, auprès des patients hospitalisés dans le service de gériatrie à orientation cardiologique et neurologique de l'hôpital Charles Foix. Si la personne répondait aux critères d'inclusion et non pas à ceux de non inclusion, la lettre d'informations (voir annexe E) lui était remise et expliquée par le médecin hospitalier en charge de ses soins. De plus, le médecin a répondu à l'ensemble des questions du sujet ayant attrait à l'étude.

De même, les soignants ont été recrutés auprès du personnel infirmier hospitalier du service de gériatrie à orientation cardiologique et neurologique de l'hôpital Charles Foix. Si le soignant répondait aux critères d'inclusion et de non inclusion, la lettre d'informations (voir annexe F) lui était remise et expliquée par le médecin hospitalier. Dans la continuité, le médecin a répondu à l'ensemble des questions du soignant ayant attrait à l'étude.

4.3.1 Protocole

Une fois le recrutement des participants effectué, avant de démarrer la capture de données, l'ensemble des soignants inclus dans l'étude ont reçu une formation et des informations sur le système automatique intégré dans un chariot médical. Lors de cet entretien l'objectif de l'étude a été rappelé et les différents composants du système et leurs fonctionnalités ont

³ Encore une fois, nous n'étudions pas les comportements observés en fonction du genre ou d'un trait humain particulier, nous n'avons donc pas enregistré d'informations spécifiques comme le genre ou la raison de leur hospitalisation

été présentés. Enfin le positionnement du chariot pour une capture vidéo et audio optimale durant la prise de sang a été défini.

Les enregistrements lors du soin se sont déroulés comme illustré dans le schéma 4.7 :

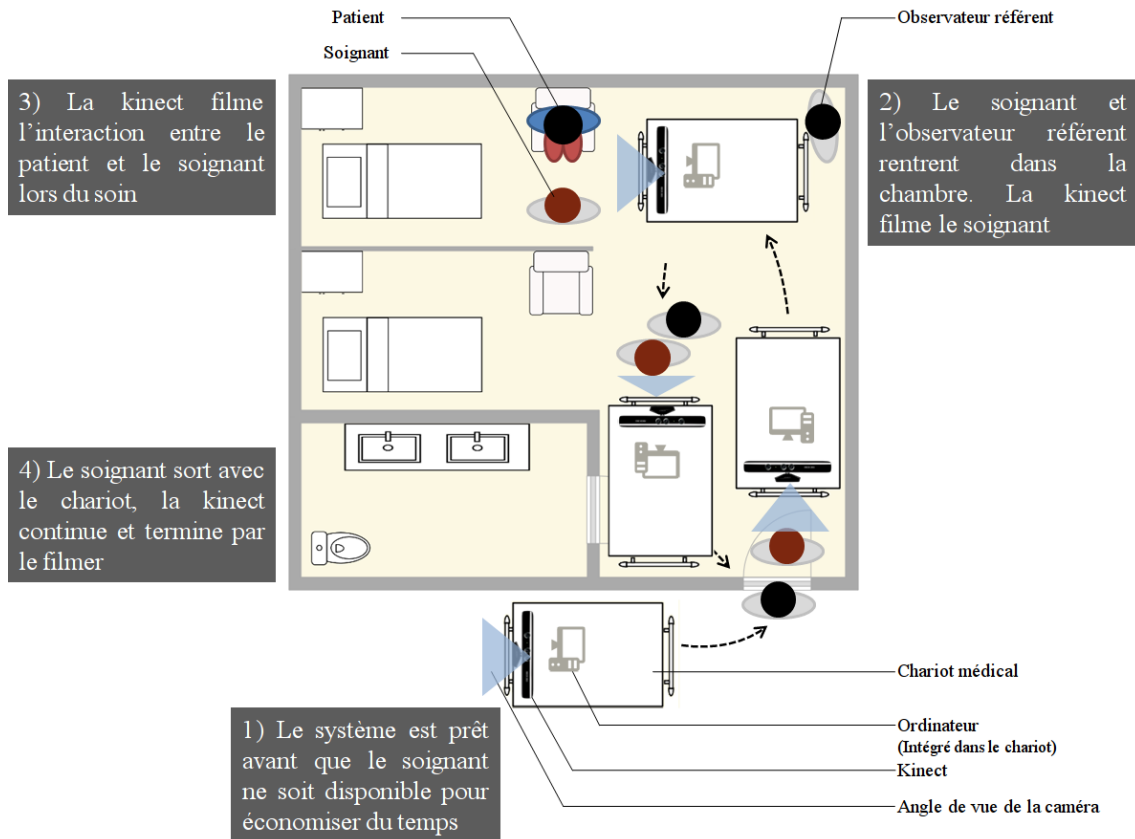


FIGURE 4.7 – 1) Le chariot médical comprenant le système embarqué est préparé en avance ; 2) Le soignant entre dans la chambre, la caméra le filme durant la première interaction avec le patient 3) le soignant positionne le chariot de manière à ce que la caméra puisse enregistrer la prise de sang 4) Le soignant sort avec le chariot et quitte la chambre. OR représente l'Observateur référent dans la chambre tout au long de l'enregistrement. Il est chargé d'évaluer l'interaction Patient/Soignant.

Le soignant entre dans la chambre du patient avec le chariot médical équipé du système embarqué, il est suivi par l'OR. Le premier contact entre le soignant et le patient est enregistré à l'aide de la Kinect filmant le soignant. Une fois arrivé à proximité du patient, le soignant positionne le chariot de manière à ce que la caméra puisse enregistrer l'ensemble de la scène d'interaction entre le soignant et le patient. Le soignant réalise le (ou les) soin(s) nécessaire(s) à la prise en charge du jour du patient, prescrite par le médecin.

Une fois l'acte terminé, le soignant sort de la chambre en poussant le chariot, la caméra continuant à le filmer. Une fois complètement sorti de la chambre avec le chariot, l'enregistrement était arrêté. Enfin, une question sur l'impact de l'utilisation du chariot et du système sur le travail de l'infirmier lui a été posée. A travers cette question nous souhaitons répondre à l'hypothèse selon laquelle la présence de caméras lors d'un soin n'impacte pas le travail des soignants ni n'altère le comportement des patients et soignants. Ainsi, une dernière question a été posée au patient ainsi qu'au soignant sur l'attention qu'ils ont porté à la caméra.

Durant la réalisation de l'interaction, l'OR a effectué la cotation des échelles ECPAI et

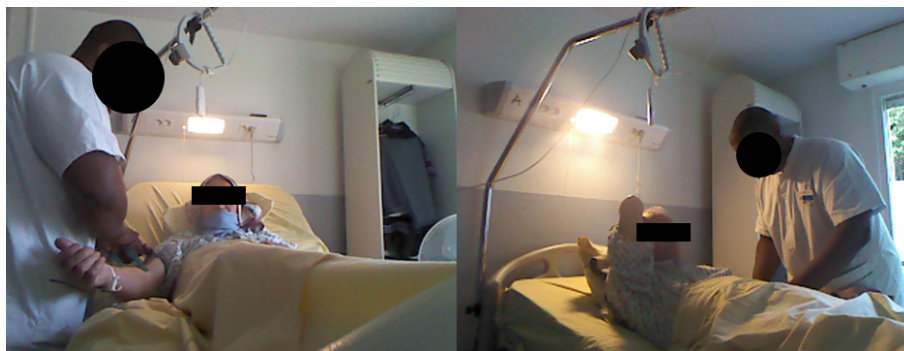


FIGURE 4.8 – Prise de vue du système lors d’un soin. Dans l’image de gauche, le patient est sous les draps. Dans celle de droite, la luminosité et le positionnement du patient n’ont pas permis de réaliser les mêmes traitement d’information que dans l’étude précédente faite dans un environnement contrôlé.

ECIIPA.

Dans la seconde partie de l’étude, L’analyse, par les OV et l’OS a été faite. Durant leur session d’évaluation chaque observateur pouvait manipuler les enregistrements et les informations à leur guise (retour en arrière, arrêt, etc.). Chaque observateur a été amené à analyser l’intégralité des vidéos à la suite durant la même session.

4.3.2 Problématiques environnementales et adaptation du protocole initial

Sur les 8 interactions enregistrées, seules 6 ont pu être exploitées. Deux enregistrements réalisés ont été supprimés suite à une qualité d’image et de son médiocre. En effet, de nombreux problèmes liés à l’environnement sont apparus lors de cette étude. La prise de sang se faisant le matin sur la période de 7h à 8h30, les conditions de luminosité étaient changeantes. L’ambiance sonore d’une chambre d’hôpital et la faible voix des participants n’ont pas permis d’avoir une qualité suffisante pour appliquer les algorithmes liés au son. De plus, les patients n’ont pas pu être mis au fauteuil afin de réaliser le soin. En conséquence ils étaient au lit sous des draps rendant les algorithmes de détection du squelette inefficaces (voir la prise de vue 4.8).

En conséquence toutes les analyses qu’ont permis les enregistrements concernaient le visage, la détection du sourire chez le patient ou le soignant.

Suite à ces problématiques, nous avons modifié le protocole pour cette étude. En revanche, nous préparons une nouvelle opération similaire à celle prévue initialement. D’autres capteurs et algorithmes seront employés. Malgré l’impact sur le temps du soin, la mise au fauteuil devient un critère d’inclusion pour le patient et est ajoutée au protocole.

Initialement, l’OS ne devait avoir accès qu’à des informations sémantiques renseignées par le système. Ces informations sémantiques sont d’une forme similaire à celles obtenues dans la figure 3.2. Afin de terminer cette étude, l’évaluation de l’OS s’est faite en utilisant les quelques informations extraites par le système et les enregistrements sonores. Ces derniers, ayant un niveau sonore trop faible et étant très bruités, ne pouvaient servir pour le système mais restent audibles pour un humain équipé d’un casque audio et installé dans un environnement calme lors de sa session. Durant les sessions d’enregistrement réalisées, très peu d’échanges verbaux ont eu lieu entre les patients et le personnel médical. Suite à l’écoute de ces enregistrements, nous avons jugés qu’aucune information personnelle ou

sensible n'était en contradiction avec nos règles éthiques établies précédemment.

Cette modification du protocole ne rentre pas en conflit avec nos objectifs initiaux. Comme dit en introduction de ce chapitre, nous ne cherchons pas à remplacer l'être humain pour remplir les échelles ECPAI et ECIIPA. En revanche, nous proposons une nouvelle approche pour réaliser cette tâche en enregistrant l'interaction et en extrayant des informations complémentaires pouvant servir à l'observateur humain. Les aléas techniques et environnementales justifient la collaboration entre l'humain et la plate-forme pour l'évaluation de l'interaction.

4.3.3 Analyse des résultats

De la même manière dont nous avons comparé les résultats de l'étude réalisée à l'INSEAD Behavioural Lab (voir la sous section 4.2.3), nous comparons ici les évaluations faites par chaque observateur. Dans un premier temps nous cherchons à déterminer si le point de vue du système permet de réaliser cette tâche d'évaluation par un observateur humain. Nous comparons donc les échelles de l'observateur référent avec celles des observateurs vidéos. Au préalable, une première étude comparative a été menée afin de vérifier que les observateurs vidéos n'ont pas analysé les comportements de manière hétérogène. Enfin, nous avons comparé les évaluations de l'observateur utilisant le système. Selon les mêmes critères définis dans la sous section 4.2.3, nous considérons qu'il y a un accord sur l'évaluation d'un comportement si l'écart entre les cases annotées est inférieur ou égal à un. S'il est supérieur ou non renseigné, on considère un désaccord.

Comparaison des évaluations entre les observateurs vidéos

Ces enregistrements ont été utilisés tels quels par les observateurs vidéo afin d'évaluer le comportement du patient et du soignant. Si un des deux OV a estimé que la position de la caméra n'était pas assez optimale pour bien observer les interactions, les deux observateurs ont pu évaluer l'intégralité de l'interaction. Une remarque commune, suite à un court entretien réalisé après leur session, a permis de faire ressortir l'intérêt d'utiliser des enregistrements plutôt que d'être présent dans la chambre lors du soin. Ceci, notamment pour une question de compatibilité d'emploi du temps de l'observateur quant à la date et l'horaire du soin, ainsi que la possibilité interrompre la lecture de l'enregistrement pour de la prise de notes ou pour remplir les échelles.

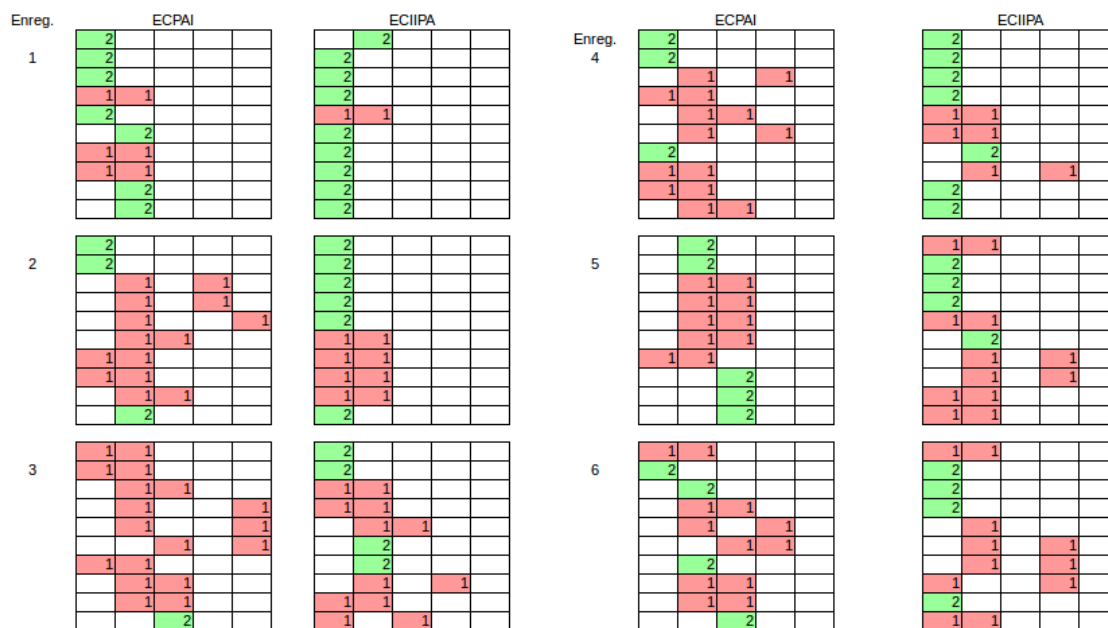


FIGURE 4.9 – Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d’observateurs évaluant le comportement de la même manière : Lorsque les deux observateurs ont évalué de manière similaire, la case correspondante dans l’échelle est représentée en vert, en rouge sinon

Nombre de cases d’écart entre les évaluations	0	1	2	3	4
Echelle ECIIPA (%)	56	28	12	2	2
Echelle ECPAI (%)	38	46	10	5	1

FIGURE 4.10 – Résultats de la comparaison des échelles annotées par les observateurs vidéos

Le tableau 4.10 indique les résultats de la comparaison entre les évaluations des deux observateurs vidéos. Nous pouvons remarquer le taux d'accords entre ces deux observateurs est égal (84%) sur les deux échelles ECIIPA, ECPAI. Suite aux conclusions de la section précédente, il n'est pas surprenant de retrouver cependant un accord parfait (écart entre les case de 0) largement supérieur lorsqu'il s'agit de l'échelle ECIIPA. Les principaux désaccords des échelles ECPAI apparaissent sur les interactions 2 et 3 L'interaction 6 est celle entraînant le plus de désaccords concernant l'échelle ECIIPA. Ces résultats démontrent que les observateurs vidéos ont été en mesure d'évaluer les comportements de manière similaire en employant une méthodologie visant à n'utiliser que les mêmes enregistrements utilisés plus tard par le système automatique.

Comparaison des évaluations entre les observateurs vidéos et l'observateur référent

Dans cette section, nous comparons les évaluations faites par l'observateur référent et les observateurs vidéos. Le tableau 4.12 est un résumé des échelles comparées des différents observateurs, montrées dans la figure 4.11.

Les taux d'accords ont été fait pour tous les observateurs vidéos en même temps. Ainsi, le taux de 25% d'écart d'une case correspond à l'écart d'une case entre l'évaluation de l'OR et celle d'un OV. Cette méthode de comparaison nous permet, dans un premier temps, de déterminer si un taux de désaccords important est obtenu. Si tel est le cas, une comparaison des évaluations deux à deux entre chaque OV avec l'OR sera nécessaire pour discriminer l'observateur générant ces désaccords. De plus, comme vu dans la section 4.3.3, il n'y a que peu de disparités entre les OVs. Ainsi si nous observons des taux indiquant de grands désaccords dans les résultats reportés dans le tableau 4.12, ils seront imputables à l'OR. A la vue du tableau 4.12, les résultats démontrent que l'OR s'accordent avec les OV. Ils démontrent que les deux méthodologies évaluations (directes et télévisuelles) n'influent pas sur les évaluations. De plus, le point de vue de la caméra est tout à fait satisfaisant pour permettre l'annotation de grilles par les OV.

En conséquence, si l'observateur système rencontre des difficultés à remplir les échelles, le point de vue de la caméra ne peut être remis en cause. Reste que les informations ne sont pas exploitables ou que la méthodologie télévisuelle ne conviennent à l'OS.

Comparaison des évaluations entre l'observateur référent et l'observateur système

Dans cette section, nous comparons les évaluations faites par l'observateur référent et l'observateur système. Le tableau 4.14 est un résumé des échelles comparées des différents observateurs, montrées dans la figure 4.13. N'ayant pas d'écart de 4 cases, cette colonne du tableau est remplacée par la colonne "Ne Se Prononce Pas". En effet, il est rappelé que l'OS n'a accès qu'aux enregistrements sonores et à quelques informations concernant le sourire et le calme du soignant. En conséquence, les échelles remplies par l'OS sont incomplètes. Ce dernier a tout de même rempli 78% des échelles ECPAI et 87% des échelles ECIIPA.

Malgré 13% de comportements non évalués, l'OS a un taux d'accord avec l'OR pour l'échelle ECIIPA de 80% dont 55% d'accords directs. Ce taux "tombe" à 70% d'accords pour l'échelle ECPAI malgré une absence de réponses pour 22% de l'échelle. Encore une fois, il est montré par cette étude que remplir les échelles ECIIPA est facilité par le comportement, semble-t-il professionnellement positif des soignants et ce, même en remplaçant la perception visuelle de l'observateur par des informations extraites par le système. Ces résultats permettent d'affirmer que l'emploi de la plate-forme comme système automatique d'aide à l'analyse et à l'évaluation du comportement du patient et du soignant lors d'un soin est satisfaisante.

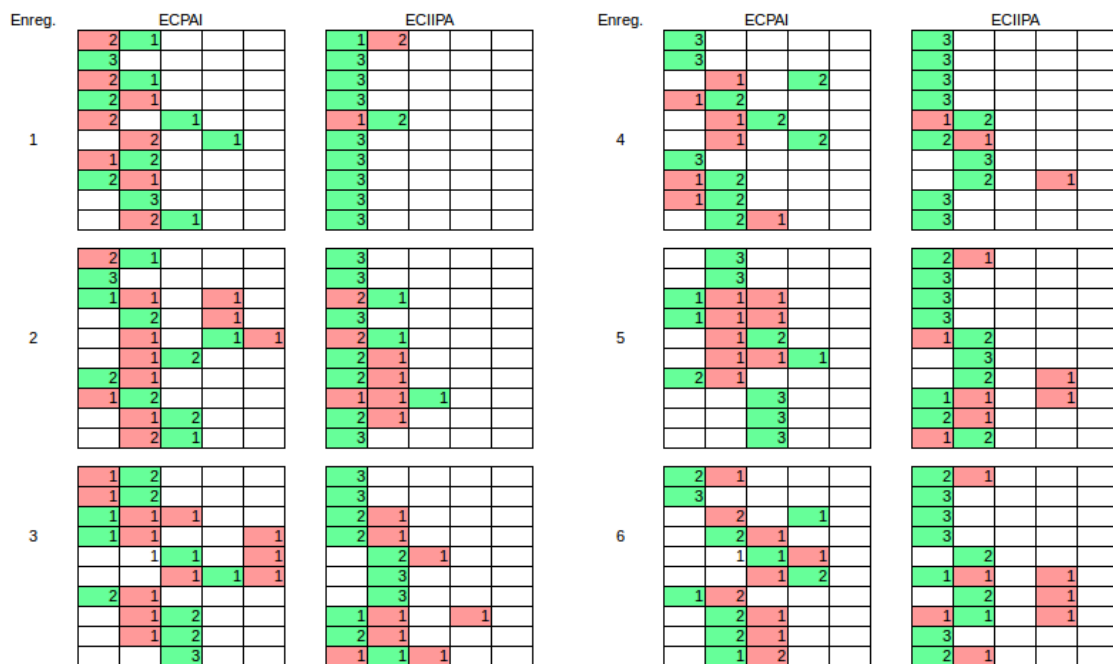


FIGURE 4.11 – Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d’observateurs évaluant le comportement de la même manière : Les cases vertes correspondent à l’évaluation de l’observateur référent, les cases rouges sont les évaluations discordantes faites par les observateurs vidéos

Nombre de cases d’écart entre les évaluations	0	1	2	3	4
Echelle ECIIPA (%)	67	25	5	2	1
Echelle ECPAI (%)	46	39	13	1	1

FIGURE 4.12 – Résultats de la comparaison des échelles annotées par les observateurs vidéos et l’observateur référent

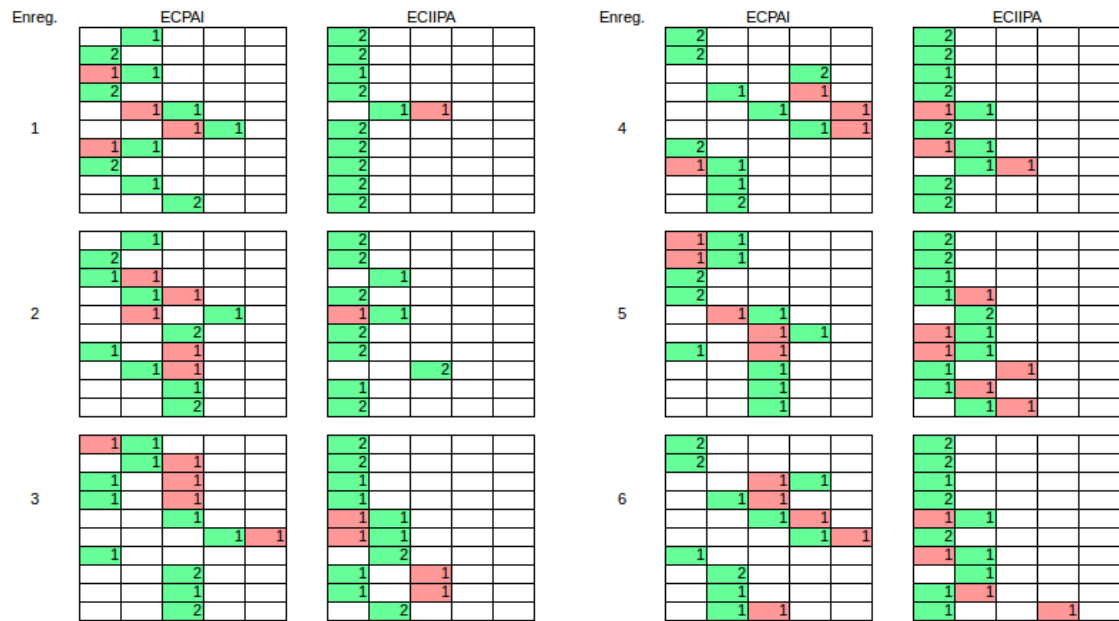


FIGURE 4.13 – Représentation simplifiée des échelles ECPAI et ECIIPA annotées par les observateurs vidéos. La valeur dans chaque case représente le nombre d’observateurs évaluant le comportement de la même manière : Les cases vertes correspondent à l’évaluation de l’observateur référent, les cases rouges sont les évaluations discordantes faites par l’observateur système

Nombre de cases d’écart entre les évaluations	0	1	2	3	4	NSPP
Echelle ECIIPA (%)	55	25	5	2	0	13
Echelle ECPAI (%)	33	35	8	2	0	22

FIGURE 4.14 – Résultats de la comparaison des échelles annotées par l’OR et l’OS

4.3.4 Conclusion

Dans ce chapitre, nous proposons une nouvelle méthodologie d'évaluation du comportement du patient et du soignant lors d'un soin, traditionnellement réalisée par un observateur présent dans la chambre. A l'aide des échelles ECPAI et ECIIPA (voir respectivement les annexes A et B) cet observateur peut annoter les comportements observés et analysés et ainsi rendre état de l'interaction passée.

Notre hypothèse de recherche est que l'analyse de l'interaction entre le patient et le soignant peut être réalisée a posteriori du soin à partir d'enregistrements. De plus, nous posons une seconde hypothèse concernant l'utilisation de ces enregistrements par la plate-forme détaillée dans le chapitre 2. Cette dernière peut extraire des informations sur les comportements des participants telles que, une fois partagées à un observateur humain, ce dernier soit en mesure de d'évaluer les comportements et de remplir les échelles ECPAI et ECIIPA.

A l'heure actuelle, il n'existe pas de corpus d'enregistrements permettant de confirmer ces hypothèses. Nous avons donc réalisé deux séries d'enregistrements d'une interaction entre un patient et un soignant lors d'un soin. La première série se déroulait dans un environnement contrôlé et était interprétée par des acteurs suivant des consignes. La seconde a été réalisée au sein de l'hôpital Charles Foix.

A la suite de la création de ces corpus, nous avons cherché à valider la première hypothèse. Les résultats, obtenus lors de la comparaison entre les échelles d'un observateur référent présent lors du soin et des observateurs utilisant des enregistrements vidéos, démontrent que cette évaluation est possible dans des conditions plus favorables : Possibilité d'arrêter la vidéo pour de la prise de notes, de revenir en arrière, etc.

Nous avons ensuite cherché à valider notre seconde hypothèse. Pour cela deux nouvelles étapes ont été mises en place. En utilisant les données des enregistrements réalisés dans un environnement contrôlé, nous avons mis en place un système d'analyse et d'évaluation automatique des comportements. Puis nous avons comparé ses évaluations avec ceux de l'observateur référent. Les résultats étant similaire nous avons donc réalisé les mêmes évaluations sur les données enregistrées en situation réelle à l'hôpital.

Les enregistrements étant de moins bonnes qualités suite à des conditions environnementales plus contraignantes, une partie des algorithmes du système automatique n'a pu être utilisée. Un des objectifs initiaux de cette thèse est de faire collaborer notre plate-forme avec l'humain pour observer le comportement humain (voir chapitre 3). Nous avons donc modifié notre protocole afin de faire collaborer le système automatique avec l'observateur de manière à ce que ses données extraites, complétées par les enregistrements sonores de l'interaction, soient utilisées par ce dernier pour évaluer le comportement.

Malgré une absence partielle d'évaluations les échelles ainsi obtenues ont été comparées à celles de l'observateur référent. Les résultats étant satisfaisants (en moyenne 75% d'accords sur les deux échelles entre les deux observateurs), nous avons validé notre seconde hypothèse.

Ces résultats gagneraient à être utilisés et comparés à ceux d'une étude complémentaire incluant plus de patients et de soignants permettant d'augmenter le nombre de d'enregistrements. De plus, ceux réalisés au sein de l'hôpital ont été faits le matin tôt. Réaliser ces enregistrements plus tard permettrait d'éviter d'évaluer des patients réveillés pour la prise de sang et donc peu enclin à interagir avec le soignant. Les comportements observables seraient alors plus diversifiés.

Chapitre 5

Conclusions et perspectives

Depuis quelques années se développent les concepts de reconnaissance automatique d'activités de la vie courante et reconnaissance automatique de comportements. Ces concepts s'articulent autour de domaines de recherche comme le "bien-vieillir à domicile" ou plus généralement des sciences du comportement humain. Les technologies alors mises en place utilisent un ensemble de capteurs positionnés de manière fixe dans des salles ou embarqués sur des robots mobiles pouvant interagir, ou non, avec l'être humain. Un système informatique récolte et analyse ensuite les données afin de déduire une activité ou un comportement correspondant aux modèles définis au sein du système.

Cette thèse, se situe dans les domaines de l'analyse automatique des activités de la vie quotidienne et l'évaluation du comportement dans le cadre d'une interaction entre humains. Nous avons mis en place une plate-forme d'évaluation du comportement humain collaborative avec des observateurs extérieurs. Elle s'appuie sur une méthodologie axée autour d'ontologies pour fusionner des informations de haut niveau au sein du système ambiant et du robot. Cette plate-forme observe le comportement et l'analyse. Les informations extraites sont ensuite transmises à un observateur extérieur chargé d'évaluer le comportement en s'appuyant sur ces données (voir figure 1.3).

Dans le premier chapitre, nous avons répondu à un ensemble de problématiques liées à l'emploi de systèmes complexes indépendants composés de plusieurs capteurs multimodaux au sein d'une même plate-forme :

Nous avons souhaité que cette dernière soit composée d'un système ambiant ubiquitaire et d'un robot mobile. Chaque système a été conçu de manière à être modulable et interopérable. Ainsi, chaque système peut se voir ajouter ou retirer un capteur sans entraîner de dysfonctionnement. Ajouter des capteurs permet de récupérer plus d'informations et donc compléter la connaissance de l'environnement de ces systèmes tandis qu'en retirer, élimine des connaissances de l'environnement. Ainsi le système fonctionnera dans un mode dégradé mais continuera à être utilisable. L'utilisation de technologies telles que ROS ou des ontologies pour la fusion de données facilite cette modularité. De plus, ces dernières facilitent l'interopérabilité en permettant la fusion d'informations multimodales de haut niveau. Ainsi, peu importe la source une information peut être comparée à une autre ou la compléter.

La mobilité du robot permet de compléter la position stationnaire des capteurs du système ambiant et, à l'inverse, l'ubiquité du système ambiant permet de compléter le point de vue limité du robot. Dans le chapitre 2, nous avons détaillé une procédure de collaboration inter-système s'appuyant sur l'interopérabilité des informations de chaque système. Toujours dans le chapitre 2, un scénario de collaboration a été présenté. Dans ce scénario, l'accomplissement d'une tâche ne pouvait se faire que si chaque système partageait avec succès ses connaissances avec l'autre. Enfin, il est à préciser que l'indépendance de chaque système n'est pas remise en cause puisque dans le chapitre 4, une version simplifiée de la

plate-forme n'employant qu'une partie du système ambiant est exploitée.

Dans la suite de la thèse, nous nous sommes intéressés à la capacité qu'a cette plate-forme à collaborer avec un être humain pour reconnaître et séquencer les activités de la vie quotidienne réalisée par un participant. Lors d'une expérience, le sujet devait réaliser une série d'activités comprenant, entre autre l'activité *dormir*, *manger* et *s'habiller*. La plate-forme complète (composée du système ambiant et du système robotique) a récupéré un ensemble de connaissances permettant d'indiquer clairement le début et la fin d'activités ou alors uniquement certaines informations spécifiques. Les ontologies, en plus de leurs rôles définies précédemment, ont permis de transmettre les connaissances de la plate-forme à un groupe d'observateurs chargés de retracer l'emploi du temps du sujet. Les résultats obtenus indiquent qu'avec un minimum d'informations, les observateurs pouvaient séquencer correctement les activités mettant donc en évidence la capacité qu'a la plate-forme à collaborer avec un humain afin de réaliser une tâche.

Dans le chapitre 4 de cette thèse, la plate-forme a été adaptée afin de lui permettre d'observer et analyser des comportements d'un patient et d'un soignant lors d'un soin. Les informations déduites de ces traitements étaient ensuite communiquées à un observateur. Cet observateur, appelé observateur système, était ensuite chargé d'évaluer le comportement du patient et du soignant à l'aide des échelles d'évaluation du comportement ECPAI et ECIIPA. Ces échelles sont traditionnellement remplies par un observateur présent lors du soin. Nous avons donc souhaité déporter l'analyse à posteriori du soin. Aucun corpus d'enregistrements d'activités et d'interactions entre un patient et un soignant existait. Nous avons donc décomposé en deux étapes distinctes la réalisation d'un corpus nous permettant de valider le fonctionnement de la collaboration entre l'observateur système et la plate-forme :

Lors de la première expérience, nous avons recruté des acteurs afin de jouer le rôle du patient et du soignant. Le recrutement des patients incluait toute personne souhaitant participer à l'étude. Celui des soignants incluait uniquement des personnes appartenant ou ayant appartenu au corps médical et ce, afin de maintenir une cohérence dans les propos et dans l'acte de la pose d'un bandage représentant le soin.

Outre des informations sur l'auto évaluation du comportement faite par les acteurs, nous avons démontré qu'une analyse utilisant les enregistrements bruts et sans traitement de la plate-forme permettait d'annoter le comportement du patient et du soignant. De plus, dans cet environnement contrôlé, la plate-forme a été en mesure de compléter, d'une manière similaire à l'observateur utilisant les enregistrements bruts, une partie des échelles ECPAI et ECIIPA. Les algorithmes employés ne pouvant compléter que la moitié de ces échelles. Enfin, un sondage nous indique que la présence de caméras lors de l'interaction n'a pas attiré une attention particulière des acteurs. Ceci est à nuancer avec le fait que chaque participant avait des recommandations sur son rôle à exécuter lors de l'interaction entraînant donc un détachement de l'attention de l'environnement dans lequel cette dernière avait lieu.

Dans un second temps, nous avons souhaité réaliser la même expérience que précédemment en conditions réelles, lors d'une prise de sang s'inscrivant dans le cadre de la prise en charge de patients hospitalisés. Des problèmes techniques et environnementaux nous ont empêché de réaliser un protocole identique à l'expérience précédente. En effet, les enregistrements donnant des qualités d'informations médiocres, nous avons modifié la méthodologie initiale. Ainsi, au lieu que l'observateur système n'utilisent que des informations provenant des connaissances du système pour évaluer le comportement, ce dernier bénéficiait des enregistrements sonores ainsi que les informations visuelles extraites par la plate-forme. Cette adaptation du protocole reste dans la continuité des travaux engagés

dans cette thèse. En effet, l’observateur système collabore avec la plate-forme en exploitant ses connaissances pour évaluer le comportement. Au final, ces échelles ont été comparées à celles d’observateurs utilisant les enregistrements audio et vidéo bruts ainsi qu’à celles d’un observateur référent présent dans la chambre du patient lors des enregistrements. Les conclusions de cette expérience montrent que les trois évaluations (référentes, vidéos et systèmes) donnent des résultats similaires. De plus, suite à des entretiens avec les observateurs, l’utilisation d’un système d’enregistrement de l’interaction pour l’évaluation a été recommandé par chaque observateur. Cette méthodologie permettant de prendre le temps nécessaire pour évaluer correctement l’interaction.

5.1 Perspectives

En dehors du temps alloué pour cette thèse, la dernière expérience faite à l’hôpital sera réitérée afin de créer un nouveau corpus pour mettre en place l’évaluation par l’observateur système des comportements du patient et du soignant en utilisant uniquement les connaissances de la plate-forme. Analyser les résultats de ces enregistrements s’inscrit dans la continuité directe de la thèse.

Dans cette thèse, nous nous sommes contentés de comparer les évaluations en fonction des observateurs. Une prochaine expérience pourrait être la mise en place d’un protocole à plus grande échelle permettant les évaluations en fonction de paramètres comme des pathologies médicales (Alzheimer, Parkinson, etc.), le genre ou l’âge du participant, etc. Ceci afin de pouvoir, si les résultats le permettent, de répondre avec un comportement adéquat en fonction des troubles dont souffre le patient et donc optimiser la relation entre un soignant et son patient.

Pour chaque expérience menée dans le chapitre 4, nous n’avons eu qu’une approche orientée système pour l’analyse des données. Nous avons tout d’abord validé qu’un observateur utilisant les enregistrements bruts était en mesure d’évaluer les comportements et ce, soit par retours de questionnaire (voir la section 4.2.3), soit par la comparaison de ses échelles avec celles d’un observateur référent (voir la section 4.3.3). Nous avons comparé les résultats des échelles sans rentrer dans les détails des comportements correspondants aux cases annotées. Nous savons que certains comportements sont plus faciles à interpréter que d’autres en fonction du vocabulaire employé lors de l’interaction par exemple.

Ainsi, la plate-forme employée ici pouvant enregistrer une interaction entre le patient et le soignant, chaque observateur pourrait travailler sur une base de données communes. De plus, augmenter le nombre d’évaluations différentes sur la même interaction permettrait d’employer des algorithmes de traitement de l’information nécessitant une quantité de données sensiblement supérieure à ce que nous avons déjà récolté et, ainsi, chercher à entraîner la plate-forme à reconnaître les comportements actuellement non reconnus.

Enfin, à partir de la plate-forme existante, d’autres échelles et d’autres interactions peuvent remplacer celles que nous avons employées. Une méthodologie standard pour l’enregistrement de corpus serait alors employée permettant de comparer les résultats obtenus en fonction des algorithmes implémentés ou encore en fonction des comportements évalués et échelles utilisées.

Bibliographie

- [Aigrain et al., 2016] Aigrain, J., Spodenkiewicz, M., Dubuisson, S., Detyniecki, M., Cohen, D., and Chetouani, M. (2016). Multimodal stress detection from multiple assessments. *IEEE Transactions on Affective Computing*.
- [Altmann, 1974] Altmann, J. (1974). Observational study of behavior : sampling methods. *Behaviour*, 49(3) :227–266.
- [Amigoni et al., 2005] Amigoni, F., Gatti, N., Pinciroli, C., and Roveri, M. (2005). What planner for ambient intelligence applications? *IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART A-SYSTEMS AND HUMANS*, 35 :7–21.
- [Anzalone et al., 2014] Anzalone, S., Tilmont, E., Boucenna, S., Xavier, J., Jouen, A. L., Bodeau, N., Maharatna, K., Chetouani, M., Cohen, D., Group, M. S., et al. (2014). How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3d+ time) environment during a joint attention induction task with a robot. *Research in Autism Spectrum Disorders*, 8(7) :814–826.
- [Atkeson et al., 2000] Atkeson, C., Hale, J., Pollick, F., Riley, M., Kotosaka, S., Schaul, S., Shibata, T., Tevatia, G., Ude, A., Vijayakumar, S., et al. (2000). Using humanoid robots to study human behavior. *IEEE Intelligent Systems and their applications*, 15(4) :46–56.
- [Baer et al., 1968] Baer, D., Wolf, M., and Risley, T. (1968). Some current dimensions of applied behavior analysis 1. *Journal of applied behavior analysis*, 1(1) :91–97.
- [Barrish et al., 1969] Barrish, H., Saunders, M., and Wolf, M. (1969). Good behavior game : Effects of individual contingencies for group consequences on disruptive behavior in a classroom 1. *Journal of applied behavior analysis*, 2(2) :119–124.
- [Billard and Dautenhahn, 1997] Billard, A. and Dautenhahn, K. (1997). Grounding communication in situated, social robots. In *Proceedings Towards Intelligent Mobile Robots Conference, Report No. UMCS-97-9-1, Department of Computer Science, Manchester University*.
- [Bourgeois, 1990] Bourgeois, M. (1990). Enhancing conversation skills in patients with alzheimer’s disease using a prosthetic memory aid. *Journal of Applied Behavior Analysis*, 23(1) :29–42.
- [Bowman et al., 2001] Bowman, M., Lopez, A., and Tecuci, G. (2001). Ontology development for military applications. In *Proceedings of the Thirty-ninth Annual ACM Southeast Conference, Athens, GA, ACM Press : New York, NY*. Citeseer.
- [Breazeal, 2004] Breazeal, C. (2004). *Designing sociable robots*. MIT press.
- [Broekens et al., 2009] Broekens, J., Heerink, M., Rosendal, H., et al. (2009). Assistive social robots in elderly care : a review. *Gerontechnology*, 8(2) :94–103.
- [Carroll et al., 2004] Carroll, J. J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., and Wilkinson, K. (2004). Jena : implementing the semantic web recommendations. In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, pages 74–83. ACM.

- [Chan et al., 2008] Chan, M., Estève, D., Escriba, C., and Campo, E. (2008). A review of smart homes—present state and future challenges. *Computer methods and programs in biomedicine*, 91(1) :55–81.
- [Chandrasekaran et al., 1999] Chandrasekaran, B., Josephson, J. R., and Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems and their applications*, 14(1) :20–26.
- [Charlet et al., 2009] Charlet, J., Baneyx, A., Steichen, O., Alecu, I., Daniel-Le Bozec, C., Bousquet, C., and Jaulent, M. C. (2009). Utiliser et construire des ontologies en médecine. le primat de la terminologie. *Technique et science informatiques*, 28(2) :145–171.
- [Chen et al., 2012] Chen, L., Hoey, J., Nugent, C., Cook, D., and Yu, Z. (2012). Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6) :790–808.
- [Cooper, 1982] Cooper, J. (1982). Applied behavior analysis in education. *Theory into practice*, 21(2) :114–118.
- [Cummins et al., 2015] Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., and Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, 71 :10–49.
- [de la Solidarité et de la Santé web, 2018] de la Solidarité et de la Santé web, M. (2018). Autonomie des personnes âgées : campagne d’information sur les aides et les solutions. <https://www.pour-les-personnes-agees.gouv.fr/actualites/autonomie-des-personnes-agees-campagne-dinformation-sur-les-aides-et-les-solutions>. Accessed : 10-10-2018.
- [Dunkels et al., 2004] Dunkels, A., Gronvall, B., and Voigt, T. (2004). Contiki—a light-weight and flexible operating system for tiny networked sensors. In *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*, pages 455–462. IEEE.
- [Durrant-Whyte and Bailey, 2006] Durrant-Whyte, H. and Bailey, T. (2006). Simultaneous localization and mapping : part i. *IEEE robotics & automation magazine*, 13(2) :99–110.
- [Fong et al., 2003] Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3-4) :143–166.
- [France et al., 2000] France, J., Dijkstra, J., Dhanoa, M., Lopez, S., and Bannink, A. (2000). Estimating the extent of degradation of ruminant feeds from a description of their gas production profiles observed in vitro : derivation of models and other mathematical considerations. *British Journal of Nutrition*, 83(2) :143–150.
- [Gan and Harris, 2001] Gan, Q. and Harris, C. (2001). Comparison of two measurement fusion methods for kalman-filter-based multisensor data fusion. *IEEE Transactions on Aerospace and Electronic systems*, 37(1) :273–279.
- [Girard and Cohn, 2016] Girard, J. M. and Cohn, J. F. (2016). A primer on observational measurement. *Assessment*, 23(4) :404–413.
- [Hall et al., 1995] Hall, J. A., Harrigan, J. A., and Rosenthal, R. (1995). Nonverbal behavior in clinician—patient interaction. *Applied and Preventive Psychology*, 4(1) :21–37.
- [Hong et al., 2009] Hong, X., Nugent, C., Mulvenna, M., McClean, S., Scotney, B., and Devlin, S. (2009). Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing*, 5(3) :236–252.
- [Hu et al., 2014] Hu, N., Bormann, R., Zwölfer, T., and Kröse, B. (2014). Multi-user identification and efficient user approaching by fusing robot and ambient sensors. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 5299–5306. IEEE.

- [Iranzo et al., 2006] Iranzo, A., Molinuevo, J. L., Santamaría, J., Serradell, M., Martí, M. J., Valldeoriola, F., and Tolosa, E. (2006). Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder : a descriptive study. *The Lancet Neurology*, 5(7) :572–577.
- [Jamshidi, 2008a] Jamshidi, M. (2008a). Introduction to system of systems. *System of Systems Engineering. Innovations for the 21st Century*, pages 1–43.
- [Jamshidi, 2008b] Jamshidi, M. (2008b). *Systems of systems engineering : principles and applications*. CRC press.
- [Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1) :35–45.
- [Kanda et al., 2004] Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children : A field trial. *Human–Computer Interaction*, 19(1-2) :61–84.
- [Khaleghi et al., 2013] Khaleghi, B., Khamis, A., Karray, F., and Razavi, S. (2013). Multisensor data fusion : A review of the state-of-the-art. *Information fusion*, 14(1) :28–44.
- [Kuffner, 2010] Kuffner, J. (2010). Cloud-enabled humanoid robots. *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on, Nashville TN, United States, Dec.*
- [Kurtz et al., 2016] Kurtz, S., Draper, J., and Silverman, J. (2016). *Skills for communicating with patients*. CRC Press.
- [Lara et al., 2013] Lara, O., Labrador, M., et al. (2013). A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials*, 15(3) :1192–1209.
- [Lemaignan, 2013] Lemaignan, S. (2013). Grounding the interaction : knowledge management for interactive robots. *KI-Künstliche Intelligenz*, 27(2) :183–185.
- [Lemaignan et al., 2010] Lemaignan, S., Ros, R., Mösenlechner, L., Alami, R., and Beetz, M. (2010). Oro, a knowledge management platform for cognitive architectures in robotics. In *IROS*, pages 3548–3553.
- [Lord et al., 1989] Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., and Schopler, E. (1989). Autism diagnostic observation schedule : a standardized observation of communicative and social behavior. *Journal of autism and developmental disorders*.
- [Madni and Sievers, 2014] Madni, A. M. and Sievers, M. (2014). System of systems integration : Key considerations and challenges. *Systems Engineering*, 17(3) :330–347.
- [Mast, 2007] Mast, M. S. (2007). On the importance of nonverbal communication in the physician–patient interaction. *Patient education and counseling*, 67(3) :315–318.
- [McGurk and MacDonald, 1976] McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588) :746.
- [Nakadai et al., 2008] Nakadai, K., Okuno, H. G., Nakajima, H., Hasegawa, Y., and Tsujino, H. (2008). An open source software system for robot audition hark and its evaluation. In *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*, pages 561–566. IEEE.
- [Nowak, 2003] Nowak, C. (2003). On ontologies for high-level information fusion. In *Information Fusion, 2003. Proceedings of the Sixth International Conference of*, volume 1, pages 657–664. IEEE.
- [Olugbade et al., 2018] Olugbade, T., Berthouze, N., Marquardt, N., and Williams, A. (2018). Human observer and automatic assessment of movement related self-efficacy

- in chronic pain : from exercise to functional activity. *IEEE Transactions on Affective Computing*.
- [O’Neill et al., 2015] O’Neill, R., Albin, R., Storey, K., Horner, R., and Sprague, J. (2015). *Functional assessment and program development*. Nelson Education.
- [Panait and Luke, 2005] Panait, L. and Luke, S. (2005). Cooperative multi-agent learning : The state of the art. *Autonomous agents and multi-agent systems*, 11(3) :387–434.
- [Pawlikowska et al., 2012] Pawlikowska, T., Zhang, W., Griffiths, F., van Dalen, J., and van der Vleuten, C. (2012). Verbal and non-verbal behavior of doctors and patients in primary care consultations—how this relates to patient enablement. *Patient education and counseling*, 86(1) :70–76.
- [Picard et al., 1995] Picard, R. W. et al. (1995). *Affective computing*. Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology.
- [Queirós et al., 2015] Queirós, A., Silva, A., Alvarelhão, J., Rocha, N. P., and Teixeira, A. (2015). Usability, accessibility and ambient-assisted living : a systematic literature review. *Universal Access in the Information Society*, 14(1) :57–66.
- [Quigley et al., 2009] Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., and Ng, A. (2009). Ros : an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan.
- [Ramos et al., 2008] Ramos, C., Augusto, J. C., and Shapiro, D. (2008). Ambient intelligence—the next step for artificial intelligence. *IEEE Intelligent Systems*, 23(2) :15–18.
- [Rashidi and Mihailidis, 2013] Rashidi, P. and Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE journal of biomedical and health informatics*, 17(3) :579–590.
- [Riboni and Bettini, 2011] Riboni, D. and Bettini, C. (2011). Cosar : hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing*, 15(3) :271–289.
- [Roecker and McGillem, 1988] Roecker, J. A. and McGillem, C. D. (1988). Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion. *IEEE Transactions on Aerospace and Electronic Systems*, 24(4) :447–449.
- [Roley et al., 2008] Roley, S., Delany, J., Barrows, C., Honaker, D., Sava, D., Talley, V., et al. (2008). Occupational therapy practice framework : Domain and process. *American Journal of Occupational Therapy*, 62(6).
- [Roussey et al., 2011] Roussey, C., Chanet, J. P., Soullignac, V., and Bernard, S. (2011). Les ontologies en agriculture. *Ingénierie des Systèmes d’Information*, 16(3) :55–84.
- [Scassellati et al., 2012] Scassellati, B., Admoni, H., and Matarić, M. (2012). Robots for use in autism research. *Annual review of biomedical engineering*, 14 :275–294.
- [Schaal, 1999] Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in cognitive sciences*, 3(6) :233–242.
- [Scherer et al., 2013] Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., and Morency, L. (2013). Automatic behavior descriptors for psychological disorder analysis. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, pages 1–8. IEEE.
- [Shabani and Fisher, 2006] Shabani, D. and Fisher, W. (2006). Stimulus fading and differential reinforcement for the treatment of needle phobia in a youth with autism. *Journal of applied behavior analysis*, 39(4) :449–452.
- [Shafer, 1976] Shafer, G. (1976). *A mathematical theory of evidence*, volume 42. Princeton university press.

- [Sirin et al., 2007] Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., and Katz, Y. (2007). Pellet : A practical owl-dl reasoner. *Web Semantics : science, services and agents on the World Wide Web*, 5(2) :51–53.
- [Skinner, 1966] Skinner, B. (1966). The phylogeny and ontogeny of behavior. *Science*, 153(3741) :1205–1213.
- [Stocco and Thompson, 2015] Stocco, C. and Thompson, R. (2015). Contingency analysis of caregiver behavior : Implications for parent training and future directions. *Journal of applied behavior analysis*, 48(2) :417–435.
- [Stone et al., 1999] Stone, A., Bachrach, C., Jobe, J., Kurtzman, H., and Cain, V. (1999). *The science of self-report : Implications for research and practice*. Psychology Press.
- [Tapia et al., 2004] Tapia, E. M., Intille, S., and Larson, K. (2004). Activity recognition in the home using simple and ubiquitous sensors. In *International conference on pervasive computing*, pages 158–175. Springer.
- [Tsai et al., 2007] Tsai, T. C., Hsu, Y. L., Ma, A. I., King, T., and Wu, C. H. (2007). Developing a telepresence robot for interpersonal communication with the elderly in a home environment. *Telemedicine and e-Health*, 13(4) :407–424.
- [Turk, 2014] Turk, M. (2014). Multimodal interaction : A review. *Pattern Recognition Letters*, 36 :189–195.
- [Valstar et al., 2014] Valstar, M., Schuller, B., Smith, K., Almaev, T., Eyben, F., Krajewski, J., Cowie, R., and Pantic, M. (2014). Avec 2014 : 3d dimensional affect and depression recognition challenge. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, pages 3–10. ACM.
- [Valstar et al., 2013] Valstar, M., Schuller, B., Smith, K., Eyben, F., Jiang, B., Bilakhia, S., Schnieder, S., Cowie, R., and Pantic, M. (2013). Avec 2013 : the continuous audio/visual emotion and depression recognition challenge. In *Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge*, pages 3–10. ACM.
- [Van Kasteren et al., 2008] Van Kasteren, T., Noulas, A., Englebienne, G., and Kröse, B. (2008). Accurate activity recognition in a home setting. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 1–9. ACM.
- [Viola and Jones, 2001] Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE.
- [Virués-Ortega, 2010] Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood : Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical psychology review*, 30(4) :387–399.
- [Wada et al., 2005] Wada, K., Shibata, T., Saito, T., Sakamoto, K., and Tanie, K. (2005). Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 2785–2790. IEEE.
- [web, 2017] web, I. (2017). 15+ reasons to use the robot operating system (ros). <https://www.intorobotics.com/15-reasons-to-use-the-robot-operating-system-ros/>. Last access : 13-08-2018.
- [Willner et al., 1976] Willner, D., Chang, C. B., and Dunn, K. P. (1976). Kalman filter algorithms for a multi-sensor system. In *Decision and Control including the 15th Symposium on Adaptive Processes, 1976 IEEE Conference on*, volume 15, pages 570–574. IEEE.

- [Wright et al., 2008] Wright, D., Gutwirth, S., Friedewald, M., Vildjiounaite, E., and Punie, Y. (2008). *Safeguards in a world of ambient intelligence*, volume 1. Springer Science & Business Media.
- [Xia et al., 2015] Xia, L., Gori, I., Aggarwal, J. K., and Ryoo, M. (2015). Robot-centric activity recognition from first-person rgb-d videos. In *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*, pages 357–364. IEEE.
- [Yingthawornsuk et al., 2007] Yingthawornsuk, T., Keskinpala, H. K., Wilkes, D. M., Shiavi, R. G., and Salomon, R. M. (2007). Direct acoustic feature using iterative em algorithm and spectral energy for classifying suicidal speech. In *Eighth Annual Conference of the International Speech Communication Association*.

Annexe A

ECPAI

ÉVALUATION DE L'INTERACTION RÉSIDENT INTERVENANT

Échelle de Comportement de la Personne Aidée au cours d'une Interaction avec un Intervenant (ECPAI)

Nom de la personne aidée :	Nom de l'intervenant :
Prénom de la personne aidée :	Fonction de l'intervenant :

Pour chaque ligne entourez un seul nombre en choisissant le comportement qui correspond le mieux à celui de la personne aidée (PA) au cours de l'interaction avec l'intervenant (I).

Comportement A	La PA a le comportement A de façon		La PA n'a ni l'un ni l'autre	La PA a le comportement B de façon		Comportement B
	nette	légère		nette	légère	
Est attentive à l'I ou le suit des yeux	2	1	2	1	2	Ignore l'I et a les yeux fermés
Cherche à communiquer verbalement avec l'I	2	1	2	1	2	Refuse de communiquer verbalement avec l'I
Semble calme	2	1	2	1	2	Semble agitée
Semble confiante	2	1	2	1	2	Semble inquiète/méfiante
Semble détendue	2	1	2	1	2	Semble crispée/tendue
Exprime ou semble ressentir du bien-être	2	1	2	1	2	Exprime ou semble ressentir de la douleur
Aide à l'action de l'I	2	1	2	1	2	S'oppose à l'action de l'I
Exprime de la sympathie envers l'I, verbalement	2	1	2	1	2	Exprime de l'hostilité envers l'I, verbalement
Idem mais physiquement	2	1	2	1	2	Idem mais physiquement
Son comportement s'est amélioré pendant l'interaction	2	1	2	1	2	Son comportement s'est dégradé pendant l'interaction
Total A = (0 à 20)		Total I = (0 à 20)		Total B = (0 à 20)		

Total A : Score interaction positive = Total I : Score indifférence/apathie = Total B : Score interaction négative = Score global d'interaction = (20 + A - B) / 2 =

Annexe B

ECIIPA

ÉVALUATION DE L'INTERACTION RÉSIDENT-SOIGNANT

Échelle de Comportement de la L'Intervenant au cours d'une Interaction avec une Personne Aidée (ECIIPA)

Nom de la personne aidée :	Nom de l'intervenant:
Prénom de la personne aidée :	Fonction de l'intervenant :

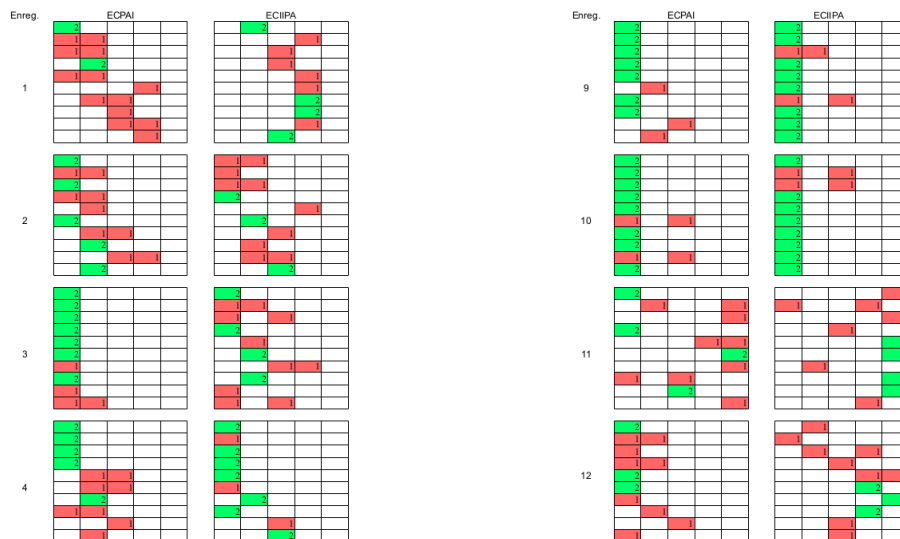
Pour chaque ligne entourez un seul nombre en choisissant le comportement qui correspond le mieux à celui de la personne aidée (PA) au cours de l'interaction avec l'intervenant (I).

Comportement A	L'I a le comportement A de façon		L'I n'a ni l'un ni l'autre	L'I a le comportement B de façon		Comportement B
	nette	légère		nette	légère	
Au départ, a cherché à capter le regard de la PA	2	1	2	1	2	Au départ, a évité le regard de la PA
Au départ, a cherché à établir un contact verbal avec la PA	2	1	2	1	2	Au départ, n'a pas cherché à établir un contact verbal avec la PA
A touché la PA après avoir établi préalablement une relation	2	1	2	1	2	A touché la PA sans avoir établi préalablement une relation
Informe la PA avant de réaliser un geste	2	1	2	1	2	Réalise un geste sans avoir préalablement informé la PA
Recherche le consentement avant de réaliser un geste	2	1	2	1	2	Réalise un geste malgré l'opposition de la PA
Semble calme et détendu	2	1	2	1	2	Semble crispé ou tendu
Semble confiant	2	1	2	1	2	Semble inquiet
Exprime ou semble ressentir du plaisir au cours de l'interaction	2	1	2	1	2	Semble être mal à l'aise ou gêné au cours de l'interaction
Est souriant et exprime de la bienveillance envers la PA	2	1	2	1	2	Exprime de l'agacement ou de la froideur envers la PA
A veillé à maintenir le contact durant toute l'interaction	2	1	2	1	2	N'a pas veillé à maintenir le contact pendant toute l'interaction
Total A = (0 à 20)			Total I = (0 à 20)			Total B = (0 à 20)

Total A : Score interaction positive = Total I : Score indifférence/apathie = Total B : Score interaction négative = Score global d'interaction = (20 + A - B) / 2 =

Annexe C

Résultats comparatifs des échelles patient/soignant de l'interaction simulée



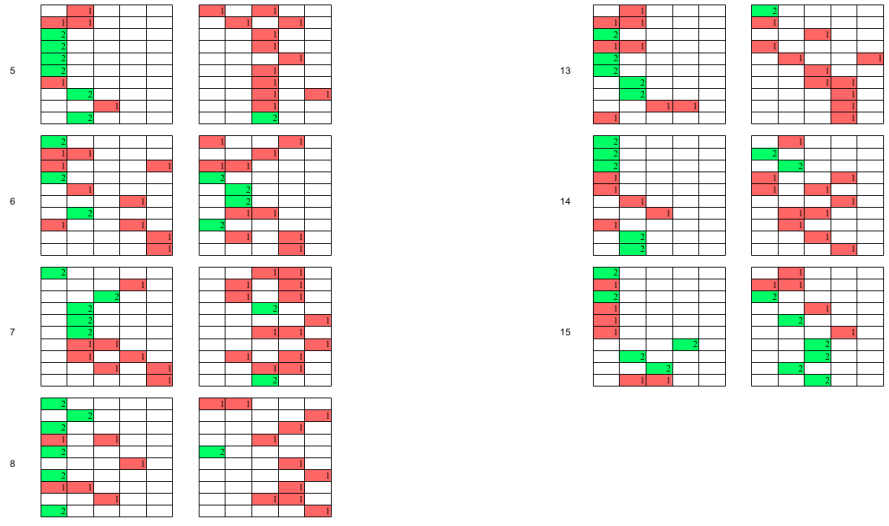
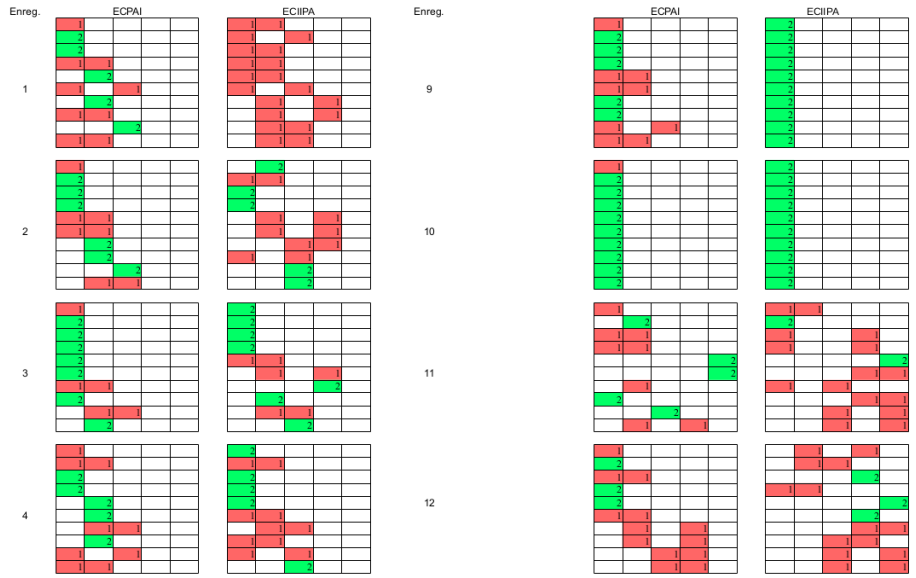


FIGURE C.1 – Ensemble des échelles complétées par les patients et les soignants. En vert, la correspondance des évaluations, en rouge la discordance



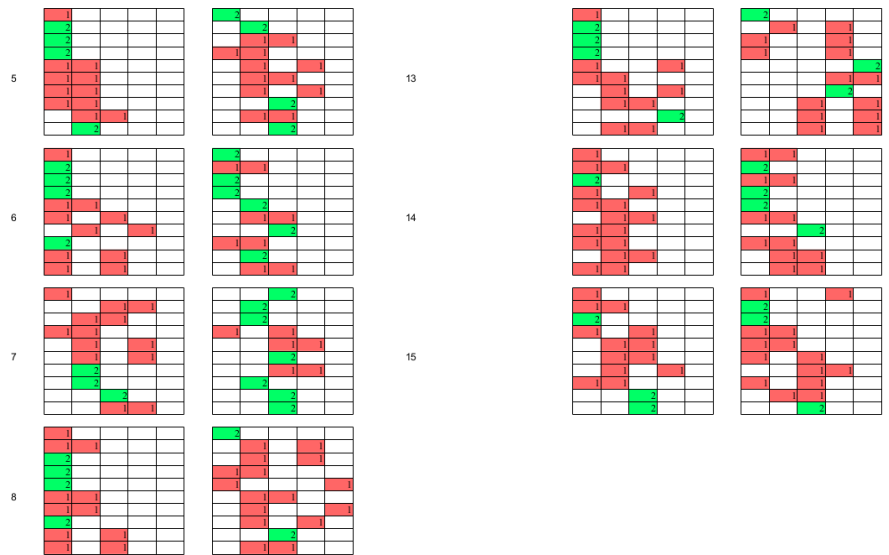
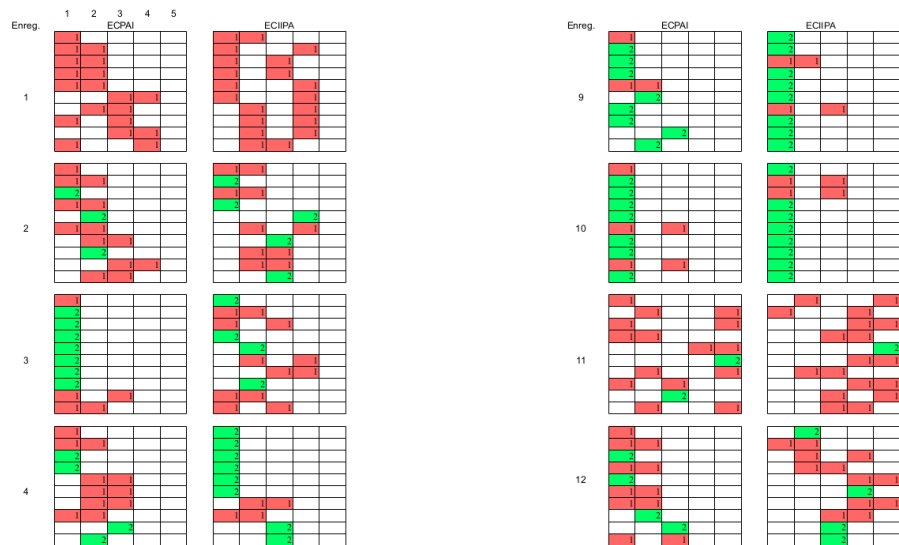


FIGURE C.2 – Ensemble des échelles complétées par les soignants et l’observateur vidéo. En vert, la correspondance des évaluations, en rouge la discordance



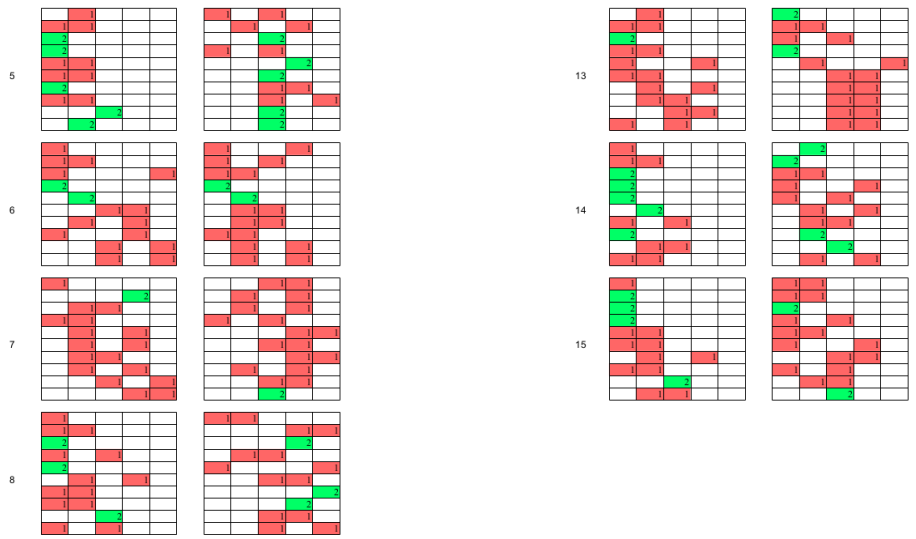
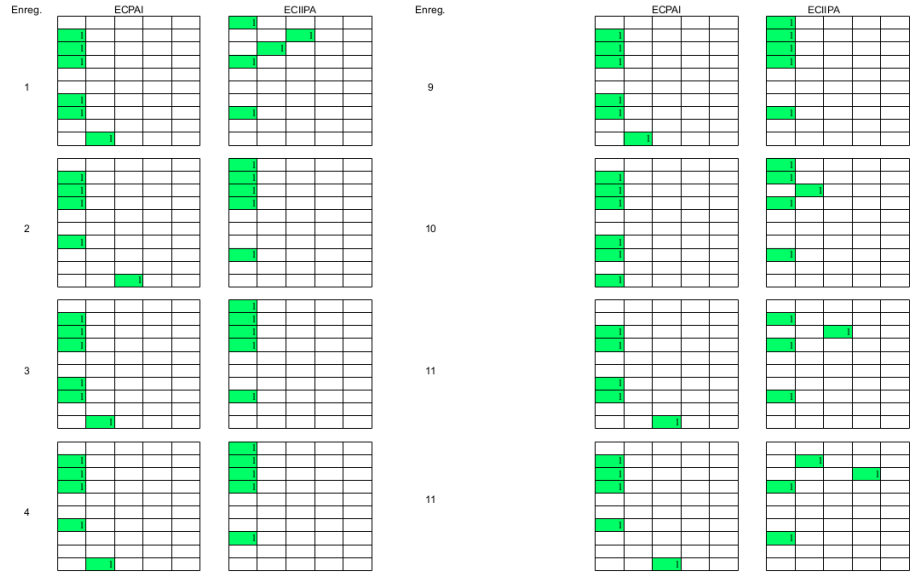


FIGURE C.3 – Ensemble des échelles complétées par les patients et l’observateur vidéo. En vert, la correspondance des évaluations, en rouge la discordance



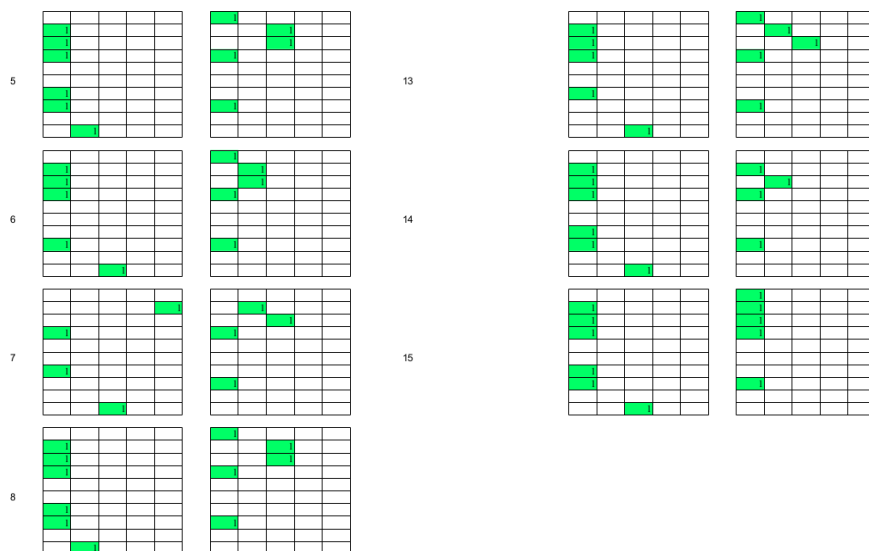
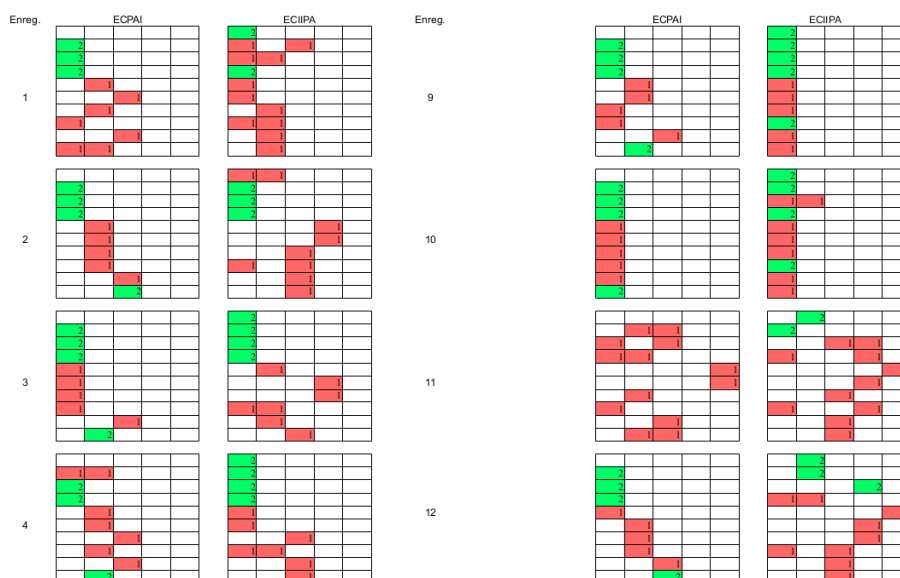


FIGURE C.4 – Ensemble des échelles complétées par le système autonome



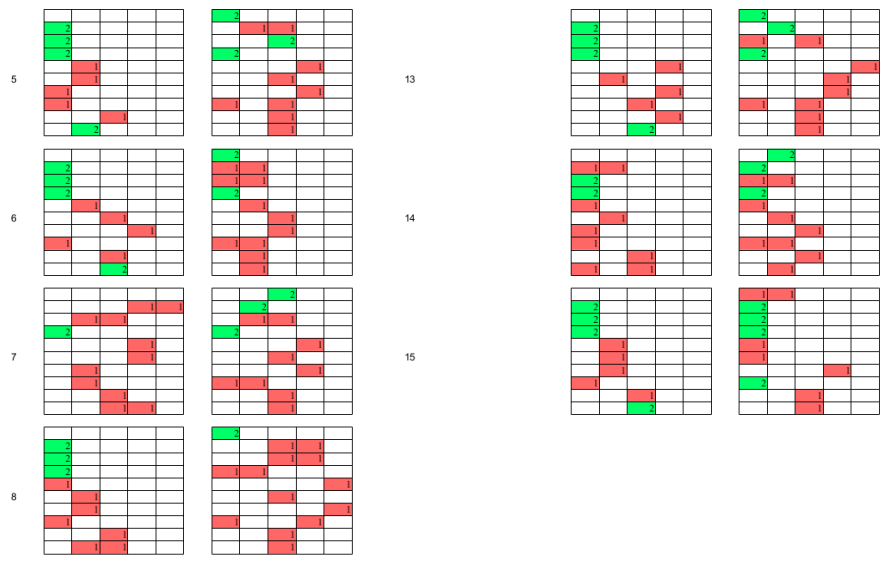


FIGURE C.5 – Ensemble des échelles complétées par le système (en vert) comparées à celles de l’observateur vidéo (en rouge)

Annexe D

Lettre de compte rendu de l'étude

D.1 Explication de ce type de recherche

L'objectif de cette expérience est d'étudier les interactions entre le patient et le soignant en combinant deux techniques subjectives (questionnaire et système automatique). De nouvelles échelles d'évaluation du comportement du patient et du soignant lors d'une intervention (ECPAI et ECIIPA) ont été développées par le professeur BELMIN (médecin, enseignant-chercheur). Afin d'annoter chaque comportement, un observateur extérieur est mandaté pour observer et rendre état de l'interaction. La présence d'un observateur durant un soin peut entraîner une modification du comportement des participants. Ainsi, faire réaliser ces observations par un système automatique permettrait des résultats objectifs avec une limitation de l'altération du comportement des personnes observées ainsi qu'un gain de temps pour le personnel médical. Le projet consiste donc à automatiser ces observations en employant un système automatique capable d'extraire des informations de l'environnement dans lequel il est placé. Afin de valider ses observations, il est important de faire une comparaison entre les données et conclusions extraites par le système avec celles du personnel médical habitué à employer ces échelles. Cette première expérience consiste donc à filmer des interactions entre deux personnes, l'un jouant le rôle de soignant et l'autre du patient. Dans un second temps, faire remplir les échelles d'évaluation de comportement par le système automatique et le personnel médical habitué à leur utilisation pour comparer les résultats.

D.2 Description plus détaillée de l'étude

Dans un premier temps, il s'agira d'employer le système automatique mis en place pour observer et enregistrer des interactions scénarisées entre un patient et un soignant. Le système s'emploiera à analyser et rendre état du comportement de chaque participant selon les échelles ECPAI et ECIIPA. Les enregistrements vidéos (images + sons) formeront un set de vidéos servant à mettre en place un questionnaire ouvert aux personnes habituées de ces échelles. Les personnes sondées auront pour objectif de donner leur avis sur les comportements observables en remplissant les mêmes échelles. Les résultats des sondés seront compilés et comparés avec ceux obtenus par le système automatique et donneront un indice de fiabilité aux données de l'ordianteur. Vous pouvez trouver plus d'informations sur ce sujet dans les ouvrages ci-dessous :

Muralidhar, S., Schmid Mast, M., & Gatica-Perez, D. (2017). How May I Help You? Behavior and Impressions in Hospitality Service Encounters. In *Proceedings of 19th ACM International Conference on Multimodal Interaction* (No. EPFL-CONF-231132).

Si vous avez des questions concernant cette étude, vous pouvez nous contacter : Emmanuel DUMONT : emmanuel.dumont@utc.fr

D.3 A propos du Centre Sorbonne Universités-INSEAD

Le Centre Sorbonne Universités-INSEAD est dédié à la recherche comportementale. Depuis son ouverture en 2002, plus de 74 000 personnes ont participé à des études au centre. La recherche menée au centre par les chercheurs de l'INSEAD, Sorbonne Universités et HECa conduit à plus de 90 publications dans des revues scientifiques, et a permis à des étudiants en thèse d'obtenir leur PhD. Les méthodes de recherche prennent la forme de questionnaires, de tâches sur ordinateur, ou d'observations de comportements en groupe.

D.4 A propos de Sorbonne Universités

58 000 étudiants, 7 685 enseignants-chercheurs, 5 200 doctorants : Sorbonne Universités rassemble des universités et des établissements d'enseignement supérieur et de recherche.

D.5 A propos de l'INSEAD

Faisant partie des plus grandes et des plus prestigieuses écoles de commerce dans le monde, l'INSEAD rassemble des personnes, des idées et des cultures du monde entier afin de changer les vies et de transformer les organisations.

D.6 Le chercheur

Emmanuel Dumont est un ingénieur en système embarqué spécialisé dans l'informatique des systèmes électroniques. Après un bref passage dans l'industrie des objets connectés, il s'intéresse au domaine de la robotique sociale en milieu médical et décide de se lancer dans une thèse. Depuis 2 ans, ses travaux portent sur les systèmes intelligents et l'analyse du comportement des personnes âgées. Les premiers devant aider les seconds à sauvegarder leur autonomie.

Annexe E

LETTRE D'INFORMATIONS ADRESSEE AUX PATIENTS POUR L'ÉTUDE

Madame, Monsieur,

Nous vous proposons de participer à une étude au sein du service de l'hôpital gériatrique Charles Foix où vous êtes hospitalisé. Cette lettre d'informations vous détaille en quoi elle consiste. Prenez le temps de lire attentivement ces informations afin de réfléchir à votre participation. N'hésitez pas à demander au médecin hospitalier, vous proposant de participer à l'étude, davantage de précisions ou de vous expliquer les éléments que vous n'avez pas compris.

E.1 BUT DE L'ETUDE

L'objectif de cette étude est d'analyser et de mesurer la qualité de l'interaction entre un patient et un soignant lors d'un acte médical classique en milieu hospitalier : la prise de sang.

D'ordinaire, cette analyse est réalisée par un observateur qualifié, présent dans la salle lors du soin et qui appartient à l'équipe médicale de l'hôpital. Dans cette étude, nous souhaitons analyser la pertinence qu'aurait l'observateur à utiliser un système automatique d'enregistrement de l'interaction soignant-patient durant le soin.

E.2 BENEFICES ATTENDUS

Nous supposons, d'une part, que l'utilisation de caméras pour filmer l'interaction réduit l'impact de la présence d'un observateur sur le comportement du patient et du soignant. D'autre part, que le traitement vidéo, réalisé à posteriori de l'enregistrement, fournira des informations supplémentaires difficilement observables en direct, ce qui permettra de compléter le point de vue de l'observateur qualifié.

E.3 DEROULEMENT DE L'ETUDE

L'étude se passe durant une prise de sang dans le cadre du soin hospitalier. Le soignant entre dans votre chambre avec le chariot médical équipé du système d'enregistrement. Une fois arrivé à proximité, le soignant positionne le chariot de manière à ce que la caméra puisse enregistrer l'ensemble de la scène d'interaction entre vous et lui. Le soignant réalise

le soin nécessaire à votre prise en charge du jour, prescrite par le médecin. Une fois l'acte terminé, le soignant sort de votre chambre avec le chariot. Pendant tout ce temps, un observateur appartenant à l'équipe médicale du service dans lequel vous êtes hospitalisé, analysera l'interaction. Il vous sera donné, ainsi qu'au soignant, un questionnaire de ressenti à remplir.

La suite de l'étude se passe après la prise de sang. Il est demandé à un second observateur d'analyser seul, la vidéo de l'interaction qui a eu lieu durant votre soin, en utilisant des échelles d'analyse du comportement fournies. La vidéo offrira en parallèle les points de vue des deux caméras. L'observateur sera amené à analyser plusieurs vidéos de soins à la suite durant la même session d'observation. Il sera averti en amont du nombre et de la durée approximative nécessaire pour réaliser toute la session. Une fois terminé, un questionnaire de ressenti à remplir est donné à l'observateur avant la fin de l'étude.

E.4 RISQUES POTENTIELS

Il n'y a aucun risque ajouté à la prise de sang habituelle. L'utilisation des technologies lors du soin n'apporte aucun risque supplémentaire puisqu'il ne vous est pas demandé de les manipuler.

E.5 FRAIS DE L'ETUDE

L'étude ne prend pas à sa charge les dépenses habituelles de vos soins et de votre accompagnement (frais médicaux, d'hospitalisation, etc.). Aucune compensation financière n'est fournie pour la participation à cette étude.

E.6 LEGISLATION - CONFIDENTIALITE

Conformément aux articles L. 1121-1 et suivants du Code de la Santé Publique, le Comité de Protection des Personnes de Strasbourg a étudié cette étude et a émis un avis favorable à sa réalisation le JJ/MM/AAAA. Toute information vous concernant recueillie pendant cet essai sera traitée de façon confidentielle. Seuls les responsables de l'étude et leur partenaire pourront avoir accès à ces données : - Le promoteur de l'étude l'association Rivages et l'investigateur de l'étude le Pr Joël Belmin 7 av de la République, 94200 IVRY SUR SEINE, tél : 01 49 59 45 54 - L'Institut des Systèmes Intelligents et de Robotique ; Université Pierre et Marie Curie - CNRS UMR 7222, - Le laboratoire de BioMécanique et BioIngénierie ; Université Technologique de Compiègne - CNRS - UMR 7338. Ainsi, les données recueillies resteront sur le territoire français.

A l'exception des personnes qui traiteront les informations dans le plus strict respect du secret médical, votre anonymat sera préservé. La publication des résultats de l'étude ne comportera aucun résultat individuel. Les données enregistrées à l'occasion de cette étude feront l'objet d'un traitement informatisé par le promoteur. S'agissant de données nominatives, vous bénéficierez à tout moment du droit d'accès et de rectification, à la portabilité, d'opposition, de réclamation, de suppression et d'oubli des données vous concernant auprès des responsables de l'étude. Ce droit est à exercer auprès de l'intermédiaire du Docteur Joël Belmin ou la Délégué à la Protection des Données de l'étude : Mme Sylvie DUPUY ; Service de la Sécurité Informatique 4, place Jussieu, 75005 PARIS.

La loi prévoit que la déclaration du fichier informatisé des données personnelles collectées pour la recherche doit être faite avant le début effectif de la recherche. Une mise en conformité des traitements avec les règles de fond du RGPD (règlement européen appliqué depuis le 25 mai 2018) a été préparée avec la CNIL. Le promoteur de la recherche et les

investigateurs s'engagent à respecter ce plan de protection des données à chaque instant en étant en mesure de démontrer la conformité des actions menées avec les exigences du règlement européen sur la protection des données personnelles (RGPD). De plus, conformément à la loi, les résultats globaux de l'étude pourront vous être communiqués si vous le souhaitez. Si vous avez des questions pendant votre participation à cette étude, vous pourrez contacter le médecin responsable de l'étude, le Dr Joël Belmin.

Vous êtes libre d'accepter ou de refuser de participer à cette étude. Cela n'influencera pas la qualité des soins qui vous seront prodigués.

Nous vous remercions d'avoir pris le temps de lire cette lettre d'information.

Annexe F

LETTRE D'INFORMATIONS ADRESSÉE AUX SOIGNANTS POUR L'ÉTUDE

Madame, Monsieur,

Nous vous proposons de participer à une étude au sein du service de l'hôpital gériatrique Charles Foix où vous travaillez. Cette lettre d'informations vous détaille en quoi elle consiste. Prenez le temps de lire attentivement ces informations afin de réfléchir à votre participation. N'hésitez pas à demander au médecin hospitalier, vous proposant de participer à l'étude, davantage de précisions ou de vous expliquer les éléments que vous n'avez pas compris.

F.1 BUT DE L'ETUDE

L'objectif de cette étude est d'analyser et de mesurer la qualité de l'interaction entre un patient et un soignant lors d'un acte médical classique en milieu hospitalier : la prise de sang.

D'ordinaire, cette analyse est réalisée par un observateur qualifié, présent dans la salle lors du soin et qui appartient à l'équipe médicale de l'hôpital. Dans cette étude, nous souhaitons analyser la pertinence qu'aurait l'observateur à utiliser un système automatique d'enregistrement de l'interaction soignant-patient durant le soin.

F.2 BENEFICES ATTENDUS

Nous supposons, d'une part, que l'utilisation de caméras pour filmer l'interaction réduit l'impact de la présence d'un observateur sur le comportement du patient et du soignant. D'autre part, que le traitement vidéo, réalisé à posteriori de l'enregistrement, fournira des informations supplémentaires difficilement observables en direct, ce qui permettra de compléter le point de vue de l'observateur qualifié.

F.3 DEROULEMENT DE L'ETUDE

L'étude se passe durant une prise de sang dans le cadre du soin hospitalier. Vous entrerez dans la chambre du patient avec le chariot médical équipé du système d'enregistrement. Une fois arrivée à proximité du patient, vous aurez à positionner le chariot de manière à ce que la caméra puisse enregistrer l'ensemble de la scène d'interaction entre vous et le patient.

Vous réaliserez le soin nécessaire à la prise en charge du jour du patient, prescrite par le médecin. Une fois l'acte terminé, vous sortirez de la chambre avec le chariot. Durant toute la durée du soin, un observateur, appartenant au service du professeur Belmin, analysera l'interaction. Il vous sera donné, ainsi qu'au patient, un questionnaire de ressenti à remplir.

La suite de l'étude se passe après la prise de sang. Il est demandé à un observateur d'analyser seul, la vidéo de l'interaction qui a eu lieu durant votre soin, en utilisant des échelles d'analyse du comportement fournies. La vidéo offrira en parallèle les points de vue des deux caméras. L'observateur sera amené à analyser plusieurs vidéos de soins à la suite durant la même session d'observation. Il sera averti en amont du nombre et de la durée approximative nécessaire pour réaliser toute la session. Une fois terminé, un questionnaire de ressenti à remplir est donné à l'observateur avant la fin de l'étude.

F.4 RISQUES POTENTIELS

Il n'y a aucun risque ajouté à la prise de sang habituelle. L'utilisation des technologies lors du soin n'apporte aucun risque supplémentaire puisqu'il ne vous est pas demandé de les manipuler.

F.5 FRAIS DE L'ETUDE

L'étude ne prend pas à sa charge les dépenses habituelles de soins et d'accompagnement du patient (frais médicaux, d'hospitalisation, etc.). Aucune compensation financière n'est fournie pour la participation à cette étude.

F.6 LEGISLATION - CONFIDENTIALITE

Conformément aux articles L. 1121-1 et suivants du Code de la Santé Publique, le Comité de Protection des Personnes de Strasbourg a étudié cette étude et a émis un avis favorable à sa réalisation le JJ/MM/AAAA. Toute information vous concernant recueillie pendant cet essai sera traitée de façon confidentielle. Seuls les responsables de l'étude et leur partenaire pourront avoir accès à ces données : - Le promoteur de l'étude l'association Rivages et l'investigateur de l'étude le Pr Joël Belmin 7 av de la République, 94200 IVRY SUR SEINE, tél : 01 49 59 45 54 - L'Institut des Systèmes Intelligents et de Robotique ; Université Pierre et Marie Curie - CNRS UMR 7222, - Le laboratoire de BioMécanique et BioIngénierie ; Université Technologique de Compiègne - CNRS - UMR 7338. Ainsi, les données recueillies resteront sur le territoire français.

A l'exception des personnes qui traiteront les informations dans le plus strict respect du secret médical, votre anonymat sera préservé. La publication des résultats de l'étude ne comportera aucun résultat individuel. Les données enregistrées à l'occasion de cette étude feront l'objet d'un traitement informatisé par le promoteur. S'agissant de données nominatives, vous bénéficierez à tout moment du droit d'accès et de rectification, à la portabilité, d'opposition, de réclamation, de suppression et d'oubli des données vous concernant auprès des responsables de l'étude. Ce droit est à exercer auprès de l'intermédiaire du Docteur Joël Belmin ou la Délégué à la Protection des Données de l'étude : Mme Sylvie DUPUY ; Service de la Sécurité Informatique 4, place jussieu, 75005 PARIS.

La loi prévoit que la déclaration du fichier informatisé des données personnelles collectées pour la recherche doit être faite avant le début effectif de la recherche. Une mise en conformité des traitements avec les règles de fond du RGPD (règlement européen appliqué depuis le 25 mai 2018) a été préparée avec la CNIL. Le promoteur de la recherche et les investigateurs s'engagent à respecter ce plan de protection des données à chaque instant

en étant en mesure de démontrer la conformité des actions menées avec les exigences du règlement européen sur la protection des données personnelles (RGPD). De plus, conformément à la loi, les résultats globaux de l'étude pourront vous être communiqués si vous le souhaitez. Si vous avez des questions pendant votre participation à cette étude, vous pourrez contacter le médecin responsable de l'étude, le Dr Joël Belmin.

Vous êtes libre d'accepter ou de refuser de participer à cette étude.

Nous vous remercions d'avoir pris le temps de lire cette lettre d'information.